.N465

D-S

# NAVAL RESEARCH LOGISTICS QUARTERLY

DEPOSITORY
JUN 24 1982

JUL 8 1982
CALIF 93940

JUL 8 1982

MARCH 1982
VOL. 29, NO. 1

OFFICE OF NAVAL RESEARCH

NAVSO P-1278

407-B

# NAVAL RESEARCH LOGISTICS QUARTERLY

## EDITORIAL BOARD

Marvin Denicoff, *Office of Naval Research*, Chairman

Murray A. Geisler, *Logistics Management Institute*

W. H. Marlow, *The George Washington University*

**Ex Officio Members**

Thomas C. Varley, *Office of Naval Research*
Program Director

Seymour M. Selig, *Office of Naval Research*
Managing Editor

## MANAGING EDITOR

Seymour M. Selig
*Office of Naval Research*
*Arlington, Virginia 22217*

## ASSOCIATE EDITORS

Frank M. Bass, *Purdue University*
Gabriel R. Bitran, *Massachusetts Institute of Technology*
Jack Borsting, *Department of Defense*
Eric Denardo, *Yale University*
Marco Fiorello, *Logistics Management Institute*
Martin J. Fischer, *Defense Communications Agency*
Saul I. Gass, *University of Maryland*
Neal D. Glassman, Office of Naval Research
Paul Gray, *Southern Methodist University*
Carl M. Harris, *Center for Management and Policy Research*
Alan J. Hoffman, *IBM Corporation*
Uday S. Karmarkar, *University of Rochester*
Paul R. Kleindorfer, University of Pennsylvania
Darwin Klingman, *University of Texas, Austin*
Kenneth O. Kortanek, *Carnegie-Mellon University*
Charles Kriebel, *Carnegie-Mellon University*
Jack Laderman, *Bronx, New York*
Gerald J. Lieberman, *Stanford University*
Clifford Marshall, *Polytechnic Institute of New York*
Richard C. Morey, *Duke University*
John A. Muckstadt, *Cornell University*
William P. Pierskalla, *University of Pennsylvania*
Thomas L. Saaty, *University of Pittsburgh*
Henry Solomon, *The George Washington University*
Wlodzimierz Szwarc, *University of Wisconsin, Milwaukee*
James G. Taylor, *Naval Postgraduate School*
Harvey M. Wagner, *The University of North Carolina*
John W. Wingate, *Naval Surface Weapons Center, White Oak*
Shelemyahu Zacks, *State University of New York at Binghamton*


The Naval Research Logistics Quarterly is devoted to the dissemination of scientific information in logistics and will publish research and expository papers, including those in certain areas of mathematics, statistics, and economics, relevant to the over-all effort to improve the efficiency and effectiveness of logistics operations.

Information for Contributors is indicated on inside back cover.


The Naval Research Logistics Quarterly is published by the Office of Naval Research in the months of March, June, September, and December and can be purchased from the Superintendent of Documents, U.S. Government Printing Office, Washington, D.C. 20402. Subscription Price: $15.00 a year in the U.S. and Canada, $18.75 elsewhere. Cost of individual issues may be obtained from the Superintendent of Documents.



The views and opinions expressed in this Journal are those of the authors and not necessarily those the Office of Naval Research.



The Naval Research Logistics Quarterly is published from appropriated funds by authority of the Office of Naval Research in accordance with NPPR P-35. Controlled circulation postage paid at Washington and additional mailing offices. Articles, letters and address changes may be forwarded to Department of the Navy, Office of Naval Research, Ballston Tower # 1. Room 607, 800 N. Quincy St., Arlington, VA. 22217.

# A THEORY OF MULTISTAGE CONTRACTUAL INCENTIVES WITH APPLICATION TO DESIGN-TO-COST

Robert W. Blanning

*Owen Graduate School of Management*
*Vanderbilt University*
*Nashville, Tennessee*

Paul R. Kleindorfer

*Department of Decision Sciences*
*The Wharton School*
*University of Pennsylvania*
*Philadelphia, Pennsylvania*

C. S. Sankar

*Computer and Information Sciences Dept.*
*Temple University*
*Philadelphia, Pennsylvania*

## 1. INTRODUCTION

A topic of increasing importance in public-sector management is the design and implementation of financial incentive systems that will encourage lower-level government units and profit-making organizations under contract to these units to use government funds efficiently and satisfy nonfinancial government objectives. One such incentive system is the Design-to-Cost (DTC) system, implemented for many major weapons acquisition projects in the Department of Defense [5], [14] in which a DTC goal is established for each project. Any deviations from the goal are corrected by changing the performance of the weapons system or by changing the number of weapons produced, or both. This paper constructs a simple model of the information, incentive and decision aspects of such an incentive system and offers insights into the tradeoffs and policy issues involved.

There is substantial literature on the theory of contracts [15], [8] and the design of incentives [2], [6], [7], [10], [9]. However, an examination of this literature discloses the need for research in two areas that are essential to an understanding of the weapons acquisition process. The first need arises whenever a development effort precedes a production effort. The dynamic incentive process—that is, a multistage process in which contractor behavior during any one stage is affected, by the incentives operative during that stage, and by an expectation of rewards or punishments in the subsequent stages—is not addressed by the literature. The second need arises due to lack of consideration of any but the most simple of hierarchies. Yet in any large government/contractor effort, the government is represented by at least three distinct organizations (the Congress, the Administration, and the Bureaucracy), and the contracting agent may be represented by several organizations as well (e.g., contractors and subcontractors). This paper addresses both the dynamic and hierarchical aspects of contracting in the DTC context (see [12], [1]).

The weapons acquisition process is viewed here as a multistage process whose characteristics change substantially over time. The acquisition process consists of (at least) three steps: (1) a development stage in which two or more contractors receive funds to design and test a prototype weapon, at the end of which a single contractor is awarded a production contract; (2) a production stage in which the winning contractor produces one or more copies of the weapon; and (3) an implementation and maintenance stage in which the weapon is maintained and modified in the field, often with some contractor support. The principal interactions occur between the first two stages (i.e., contractors behave differently during the development stage as the award of the production contract is uncertain).

Hence, we will confine our analysis to the first two stages. We assume that the contract during development stage is a fixed-price contract and the contract during production stage is an incentive contract including (1) full cost recovery, (2) a reward or penalty depending on the cost of production relative to a negotiated cost target, and (3) a reward or penalty depending on weapon performance relative to a prespecified performance target.

The hierarchical nature of the DTC system is also emphasized in our research. We examine four levels of government and contractor hierarchy. At the highest level, representing the Congress and the Administration, DTC goals and allowable probabilities of exceeding these goals are established. (Weapon cost and performance are assumed to be random variables whose mean values are controllable.) At the second level, representing DoD and the appropriate military service, the DTC goal is partitioned into two subgoals, one for the development stage and one for the production stage, and the contractors participating in the development stage are selected. At the third level, representing the military service and its project managers, most of the parameters in the incentive system are established and the production contract is awarded. (In our analysis, we assume that the decisions at this level are established by decision rules known in advance to both the government and the contractors.) At the fourth level, representing the contractors, contract parameters are negotiated and the levels of contractor effort (number of personnel, cost of raw material, etc.) are chosen for the two stages.

In the following section a model of the DTC incentive system is developed. The model is solved to determine the impact of government decisions (contractual incentive parameters, the allocation of the DTC goal between development and production stages, and the level of risk acceptable to the government) and the technological and market environment of the project on outcomes such as the quality of the weapon produced, the cost to the government, the profit of the firms in the industry, the risks assumed by the government and the risks assumed by the firms. Some of the policy implications of the model are then illustrated by a series of examples.

## 2. THE BASIC MODEL

We consider a given project and assume that Congress has established a Design-to-Cost (DTC) goal, $G$ for the project. $G$ is understood to be a constraint on total project cost and it is assumed that $G$ may be exceeded only with ex ante probability $\gamma$. One might anticipate that DoD would set $\gamma$ strategically to trade off the transactions costs of exceeding budgets and exposing itself to (re) appropriations hearings against the internal transactions costs which occur if $\gamma$ is small.

We assume that $n$ firms have been preselected as candidates for carrying out the project, in two stages. In the development stage, the $n$ firms compete against one another in producing the best design. In the second stage, the firm with the best first-stage design is awarded (the opportunity to bid on) a production contract. To state the problem precisely we need the following notation.

$e_{si}$ = Effort expended by firm $i$ in stage $s$. In the development stage, $s = d$, and in the production stage, $s = p$;

$Q_{si}(e_{si})$ = Quality (or performance level) achieved by firm $i$ in stage $s$;*

$C_{si}(e_{si})$ = Costs incurred by firm $i$ in stage $s$ as a function of effort expended.

*We assume that a single measure of quality describes adequately the performance of the system and that this measure is additive across the development and production stages (i.e., that development quality plus the quality added during the production stage equals total quality). In practice the measure is multidimensional, and many of the dimensions appropriate to the development stage are not the same as those appropriate to the production stage. The former emphasize mission requirements, proper exploitation of new technology, etc. The latter emphasize quality control, delivery schedules, etc. We assume that the government and the contractor can agree on a procedure (e.g., the calculation of a weighted sum of the various dimensions of quality) that will result in a single additive measure.

DoD is assumed to consider the following types of contract. All development contracts are firm fixed price contracts with each of the $n$ firms involved receiving $G_d/n$ dollars.* $G_d \leqslant G$ is therefore the total development cost to the government. The production contract, if awarded to firm $i$, is assumed to be a general incentive contract with payments above costs to firm $i$ specified as:

$$\tilde{\Pi}_{pi}\,(T_{pi}, e_{pi}, \tilde{Q}_d) = a\,T_{pi} + b\,[T_{pi} - \tilde{C}_{pi}(e_{pi})] + R_i(\tilde{Q}_d + \tilde{Q}_{pi}(e_{pi})), \tag{1}$$

where random quantities have a $\sim$ over them, and where

$T_{pi}$ = Target cost rate, negotiated by firm $i$ at the beginning of the production stage; $T_{pi}$ is assumed constrained to be nonnegative (negative bids are not allowed);

$Q_d$ = Cumulative progress in quality of the project during the development stage, which is the starting point for the production stage;

$a,b$ = Contract incentive parameters, where $a \geqslant 0$, $0 \leqslant b \leqslant 1$;

$R(q)$ = Performance incentive payment, for firm $i$, expressed as a function of total quality achieved over both stages.

At the end of the development stage, DoD would have spent exactly $G_d$ dollars, leaving $G_p = G - G_d$ dollars in the overall project budget. Suppose firm $i$ achieves the best performance in the development stage, i.e., suppose

$$\tilde{Q}_{di}(e_{di}) = \tilde{Q}_d = \max_{1 \leqslant j \leqslant n} \tilde{Q}_{dj}(e_{dj}). \tag{2}$$

We assume that if (2) obtains, then firm $i$ is given the exclusive right to bid on a production contract. In a realistic setting, one might assume that more than one of the leading firms at the end of the development stage is given the opportunity to bid on a production contract. This possibility is excluded here. Thus, it is assumed that the leading development firm, say $i$, is interested at the beginning of the production stage in setting $T_{pi}$, $e_{pi}$, $a$, and $b$ so as to maximize its profits in development stage $\bar{U}_{pi}(T_{pi}, e_{pi}, Q_d)$ where

$$\bar{U}_{pi}(T_{pi}, e_{pi}, Q_d) = E\{\tilde{\Pi}_{pi}\,(T_{pi}, e_{pi}, Q_d) + F_i(\tilde{Q}_{di}(e_{di}), \tilde{Q}_{pi}(e_{pi}))\,|\,\tilde{Q}_{di}(e_{di}) = Q_d\}, \tag{3}$$

where $F_i(q_d,q_p)$ represents expected follow-on benefits to firm $i$ (e.g., in terms of maintenance contracts, future benefits from the technology developed, etc.)† $\tilde{\Pi}_{pi}$ is given in (1), and $\tilde{\Pi}_{pi} + \tilde{C}_{pi}$ represents total (incentive plus cost) payments made by the government in the production stage.

---

*In general one would expect some sort of cost reimbursement to take place during the development stage. However, we assume a fixed price contract to contrast firm behavior in this stage with firm response to government-established incentives in the production stage. In Sections 5 and 6 we examine the implications of this assumption—primarily, that some firms may decline to participate in the development process. This may occur even under cost reimbursement if some firms find that their expected profits, although nonnegative, are less than the profits that would result from alternative earnings opportunities.

†It was mentioned in the introduction that the acquisition process consists not only of a development and a production stage, but also an ongoing stage of operation and maintenance, including retrofitting and occasionally major modification. We do not explicitly consider this third stage, so that we may focus more closely on the relationship between development and production. Thus, we are concerned here with acquisition costs and not with life cycle costs. However, some of this may be captured in the quality measure and the follow-on benefits. That is, the firm that produces a reliable and maintainable system is likely to develop a reputation that will lead to significant follow-on benefits. The impact of these follow-on benefits is first explained on page 4 and is discussed in more detail later in the manuscript. (See also footnote on page 7.)

Of course, firm $i$ will be subject to some constraints in indulging with its preferences as represented by (3). Indeed, we assume that "$a$" is fixed in advance by the government and the following holds for variables $(T_{pi}, e_{pi}, b)$:

$$\Pr\{\tilde{\Pi}_{pi}(T_{pi}, e_{pi}, Q_d) + \tilde{C}_{pi}(e_{pi}) \geqslant G_p\} \leqslant \gamma, \tag{4}$$

where $\gamma$ is specified by the Congress and the Administration. The fact that firms accept (4) as a constraint, of course, presumes that in the case of costs overruns, acceptable auditing practices can expose and penalize firms which cannot make a credible ex post case that (4) was observed in their planning. This dependence of contractual incentives on (legitimate) enforcement and monitoring procedures cannot be overemphasized.

Beyond fixing "$a$" and imposing (4), we will assume that production contracts are negotiated through one of two methods† (firm $i$ is the leading development firm):

Method 1, *M1*: "$b$" is fixed ex ante and any $T_{pi}, e_{pi}$ satisfying (4) will be accepted by DoD.

Method 2, *M2*: Firm $i$ and DoD negotiate $(T_{pi}, e_{pi}, b)$ at the beginning of the production stage such that (4) is satisfied and such that a Pareto efficient point is reached between firm $i$ and DoD. The preferences of firm $i$ are represented by (3). DoD is assumed to have preferences represented by a utility function $U_D(C, CO, Q)$, where $\tilde{Q} = \tilde{Q}_d + \tilde{Q}_{pi}(e_{pi})$ is final project quality, $\tilde{C} = G_d + \tilde{\Pi}_{pi} + \tilde{C}_{pi}$ is total project cost, and $\tilde{CO} = \tilde{C} - G$ is the cost overrun.

Formally, we may represent the two production stage decision processes just described as follows:

$$M1\text{: Maximize (3) with respect to } (T_{pi}, e_{pi}), \text{ subject to (4).} \tag{5}$$

$$\begin{aligned} M2\text{: } &\underset{(T_{pi}, e_{pi}, b)}{\text{Maximize}} \; [\alpha \bar{U}_{pi}(T_{pi}, e_{pi}, Q_d) + \\ &(1-\alpha)\, E\{U_D(G_d + \tilde{\Pi}_{pi} + \tilde{C}_{pi},\; G_d + \tilde{\Pi}_{pi} + \tilde{C}_{pi} - G,\; Q_d + \tilde{Q}_{pi}(e_{pi}))\}] \\ &\text{subject to (4), } T_{pi} \geqslant 0,\; e_{pi} \geqslant 0 \text{ and } 0 \leqslant b \leqslant 1, \end{aligned} \tag{5'}$$

where $\alpha$ is between 0 and 1 and reflects the relative bargaining power of the contractor against DoD, $\bar{U}_{pi}$ is defined in (3), $\tilde{\Pi}_{pi}$ is given in (1), $\tilde{C}_{pi} = \tilde{C}_{pi}(e_{pi})$ is the cost for the production stage, and $Q_d$ is the observed realization of (2). We will define the optimal solution value to (5) or (5′) as $V_{pi}(Q_d)$; this is the optimal expected return for firm $i$ if the ending quality level in (2) is $Q_d$ and firm $i$ is awarded the production contract.

Now consider the development stage. Each of the $n$ firms involved may be assumed to maximize the sum† of present benefits and expected follow-on benefits ($V_p(Q_d)$ if firm $i$ is allowed to bid on the production contract). Expected follow-on benefits may then be written:

$$\text{Expected follow-on benefits} = \begin{cases} 0 & \text{if } \tilde{Q}_{di}(e_{di}) < \tilde{Q}_d \\ V_{pi}(\tilde{Q}_d) & \text{if } \tilde{Q}_{di}(e_{di}) = \tilde{Q}_d \end{cases}. \tag{6}$$

*Method 1 was analyzed by McCall [11] for a static problem and neglecting (4). He showed the possibility of a bias in favor of inefficient firms arising from opportunity cost considerations. Such effects are largely ignored here, though briefly considered in the spirit of Canes [3] and Cummins [4], who also did not consider any constraints similar to (4).

†We ignore discounting here for notational convenience.

From (6) we see that an expected profit-maximizing contractor would solve the following problem in determining his level of effort $e_{di}$ in the development stage:

$$\text{(7)} \qquad \underset{e_{di}}{\text{Max}}\ E\{(G_d/n) - \tilde{C}_{di}(e_{di}) + V_{pi}(\tilde{Q}_{di}(e_{di}))\tilde{A}_i(e_{d1}, \ldots, e_{dn})\},$$

where $\tilde{C}_{di}(e_{di})$ is the cost incurred in stage $d$ for firm $i$ and where $\tilde{A}_i(e_{d1}, \ldots, e_{dn})$ equal to 1 if $\tilde{Q}_{di}(e_{di}) = \tilde{Q}_d = \underset{j}{\text{Max}}\{\tilde{Q}_{dj}(e_{dj})\}$ and 0 otherwise. Note that the probability that firm $i$ is allowed to bid on the production contract (i.e., Pr $\{\tilde{A}_i = 1\}$) depends on the level of effort of all the $n$ firms involved. Denote the optimal solution value in (7) by $V_{di}(\underline{e}_d, G_d, n)$, where $\underline{e}_d = (e_{d1}, \ldots, e_{dn})$.

The final step is the determination of $\underline{e}_d$. This problem may be formulated as a noncooperative game, with utility functions $V_{di}(\underline{e}_d, G_d, n)$. We are interested in a Nash solution $\underline{\hat{e}}_d = \underline{\hat{e}}_d(G_d, n)$ to this game, i.e., a joint strategy $\underline{\hat{e}}_d$ satisfying

$$\text{(8)} \qquad V_{di}(\underline{\hat{e}}_d, G_d, n) = \text{Max}\{V_{di}(\hat{e}_{d1}, \ldots, \hat{e}_{di-1}, e_{di}, \hat{e}_{di+1}, \ldots, \hat{e}_{dn}) | e_{di} \geqslant 0\},$$

for every $i \in \{1, \ldots, n\}$.*

Assuming $\underline{\hat{e}}_d(G_d, n)$ is unique (see below) for each $G_d$ and $n$, the random variables $\tilde{C}$, $\tilde{CO}$, and $\tilde{Q}$ are determined by $G_d$ and $n$ through $\underline{\hat{e}}_d$. DoD is then interested in determining $G_d$ (and possibly also $n$) so that its expected utility $E\{U_D(\tilde{C}, \tilde{CO}, \tilde{Q})\}$ is maximized. If firm $i$ is awarded the production contract, then

$$\text{(9)} \qquad \tilde{C} = \text{COST} = G_d + (\tilde{\Pi}_{pi} + \tilde{C}_{pi})$$

$$\text{(10)} \qquad CO = \text{COST OVERRUN} = \tilde{C} - G,$$

$$\text{(11)} \qquad \tilde{Q} = \text{QUALITY} = \tilde{Q}_{di} + \tilde{Q}_{pi}.$$

Thus, DoD wishes to set $G_d$ (and possibly $n$) so as to

$$\text{(12)} \qquad \underset{0 \leqslant G_d \leqslant G}{\text{Max}} \sum_{i=1}^{n} E\{U_D(G_d + \tilde{\Pi}_{pi} + \tilde{C}_{pi}, (G_d + \tilde{\Pi}_{pi} + \tilde{C}_{pi} - G), \tilde{Q}_{di} + \tilde{Q}_{pi})\} \cdot \text{Pr}\{\tilde{A}_i = 1\},$$

where all quantities are evaluated at $\underline{\hat{e}}_d(G_d, n)$, e.g.,

$$\tilde{\Pi}_{pi} = \tilde{\Pi}_{pi}\ (T_{pi}(\tilde{Q}_{di}(\hat{e}_{di})),\ \hat{e}_{pi}(\tilde{Q}_{di}(\hat{e}_{di})),\ \tilde{Q}_{di}(e_{di})),$$

where $\hat{T}_{pi}(Q_d), \hat{e}_{pi}(Q_d)$ are the optimal solution to (5)-(5′) for given $Q_d$. Major problems occur in solving (5)-(5′) and in obtaining $\hat{e}_d(G_d, n)$, to which we now turn.

## 3. SOLUTION—METHOD 1

In order to obtain analytical results, it is necessary to make assumptions about the forms of the probability distributions and reward functions. Specifically, we assume for each $i = 1, \ldots, n$ that:

1. $\hat{C}_{di}(e_{di})$ is random quantity with expected value $e_{di}^2$.

2. $\tilde{Q}_{di}(e_{di})$ is exponentially distributed, independently of $\{\tilde{Q}_{dj}(e_{dj}) | j \neq i\}$ with expected value $q_{di}e_{di}$, where $q_{di} > 0$.

*This is a Nash equilibrium, not a dominant strategy equilibrium. In other words, if all but one of the players should adopt the equilibrium strategy and the one player should depart, his profits will be reduced. However, if two or more players should collude (by departing from the Nash equilibrium and sharing their profits), they might each be able to realize higher profits. This behavior, which would almost certainly constitute conspiracy to defraud the government, is not considered in our model. In addition, we assume that each competitor for the production contract is aware of the technological capabilities and the values (e.g., for follow-on benefits) of the other competitors.

3. $\tilde{C}_{pi}(e_{pi})$ and $\tilde{Q}_{pi}(e_{pi})$ are jointly normal with respective means $e_{pi}^2$ and $q_{pi}e_{pi}(q_{pi} > 0)$, respective variances $\sigma_{pi}^2$ and $\eta_{pi}^2$, and with positive correlation coefficient $\delta_{pi}$.

4. $R_i(Q) \equiv 0$, i.e., performance incentive payments are nil.

5. $F_i(Q_d, Q_p) = H_i + h_{di}Q_d + h_{pi}Q_p$, where $H_i, h_{di} \geqslant 0$, $h_{pi} \geqslant 0$ are constants.

For this data we may write (5) as

$$\underset{T_{pi}, e_{pi}}{\text{Max}} \; [(a+b)T_{pi} - b\,e_{pi}^2 + H_i + h_{di}Q_d + h_{pi}q_{pi}e_{pi}] \tag{13}$$

subject to:

$$\Pr\{(a+b)T_{pi} - b\,\tilde{C}_{pi}(e_{pi}) + \tilde{C}_{pi}(e_{pi}) \geqslant G_p\} \leqslant \gamma. \tag{14}$$

Collecting terms, (14) may be rewritten as:

$$\Pr\{[(1-b)\,\tilde{C}_{pi}(e_{pi})] \geqslant [G_p - (a+b)\,T_{pi}]\} \leqslant \gamma. \tag{15}$$

Since $\tilde{C}_{pi}$ is normal, $(1-b)\tilde{C}_{pi}$ is also normal with mean $(1-b)e_{pi}^2$ and variance $(1-b)^2\sigma_{pi}^2$ so (15) may be expressed as

$$[(1-b)e_{pi}^2 + (a+b)\,T_{pi} - G_p] + (1-b)\sigma_{pi}K(\gamma) \leqslant 0 \tag{16}$$

where $K(\gamma)$ is the $(1-\gamma)^{\text{th}}$ fractile of the unit normal, i.e., $\Pr\{\tilde{N}(0,1) \geqslant K(\gamma)\} = \gamma$.

Define $k_{pi}(\gamma, b)$ through

$$k_{pi}(\gamma, b) = K(\gamma)(1-b)\,\sigma_{pi}. \tag{17}$$

Then (16) becomes

$$[(1-b)e_{pi}^2 + (a+b)\,T_{pi} - G_p] \leqslant -\,k_{pi}(\gamma, b). \tag{18}$$

Thus, the constraint (14) may be written as (18). Since $b \leqslant 1$, we see that (18) defines a convex region for every value of $Q_d$. Note also that $\partial k_{pi}/\partial\gamma < 0$ and $\partial k_{pi}/\partial b < 0$. Thus, as $\gamma$ or $b$ decreases the constraint region becomes larger. Similarly, as $Q_d$ decreases the constraint region becomes larger.

To find the optimal $T_{pi}$, $e_{pi}$ in (13), note that whatever $e_{pi}$ is, the optimal $T_{pi}$ will be set so that (18) holds as an equality since otherwise firm $i$ could simply increase $T_{pi}$ with consequent higher profits. Solving for $(a+b)T_{pi}$ in (18), we see therefore that, at the optimum,

$$(a+b)\,T_{pi} = G_p - k_{pi}(\gamma, b) - (1-b)e_{pi}^2. \tag{19}$$

Thus, substituting in (13) for $(a+b)T_{pi}$, the following problem characterizes the optimal $e_{pi}$.

$$\text{Max}\,[-e_{pi}^2 - k_{pi}(\gamma, b) + G_p + H_i + h_{di}Q_d + h_{pi}q_{pi}e_{pi}], \tag{20}$$

subject to $e_{pi} \geqslant 0$ and $T_{pi} \geqslant 0$. Using (19), the nonnegativity constraint on $T_{pi}$ may be expressed in terms of $e_{pi}$ as

$$(1-b)e_{pi}^2 \leqslant G_p - k_{pi}(\gamma, b). \tag{21}$$

Thus, the problem of interest is to maximize (20) subject to $e_{pi} \geqslant 0$ and (21). This simple quadratic programming problem has the solution

$$\hat{e}_{pi} = \text{Min}\left\{\left(\frac{h_{pi}q_{pi}}{2}\right), \left(\frac{G_p - k_{pi}(\gamma, b)}{1-b}\right)^{\frac{1}{2}}\right\} \tag{22}$$

and the optimal target cost $T_{pi}$ is therefore determined by (19) as

$$\hat{T}_{pi} = \frac{G_p - (1-b)\hat{e}_{pi}^2 - k_{pi}(\gamma,b)}{(a+b)}. \tag{23}$$

Finally, the optimal value of the objective function in (20) (respectively in (13)) is obtained by substituting $\hat{e}_{pi}$ for $e_{pi}$ in (20). This yields

$$V_{pi}(Q_d) = K_{pi} + h_{di}Q_d, \tag{24}$$

where $K_{pi}$ is independent of $Q_d$ and is given explicitly by

$$K_{pi} = G_p - k_{pi}(\gamma,b) + H_i - \hat{e}_{pi}^2 + h_{pi}q_{pi}\hat{e}_{pi}. \tag{25}$$

Notice from (22) that firm $i$ will expend only the minimum effort (here $\hat{e}_{pi} = 0$) in stage $P$ under Method 1 contracting unless there is some promise of follow-on rewards from such effort (i.e., unless $h_{pi} > 0$).*

From (7) and (24) we see that firm $i$ solves the following problem in determining its level of development effort $e_{di}$:

$$\underset{e_{di}\geqslant 0}{\text{Max}}\ E\{(G_d/n) - e_{di}^2 + [K_{pi} + h_{di}\tilde{Q}_{di}(e_{di})]\tilde{A}_i(e_{di},\underline{e}_d^i,n)\}, \tag{26}$$

where $\underline{e}_d^i = (e_{d1}, \ldots, e_{di-1}, e_{di+1}, \ldots, e_{dn})$. We have used the assumption in (26) that $E\{\tilde{C}_{di}(e_{di})\} = e_{di}^2$ and also the fact $\tilde{A}_i(\underline{e}_d,n) = 1$ precisely when firm $i$ achieves the maximum in (2); otherwise $\tilde{A}_i(\underline{e}_d,n) = 0$.

We first evaluate the following expression in (26):

$$EP = E\{[K_{pi} + h_{di}\tilde{Q}_{di}(e_{di})]\tilde{A}_i(e_{di},\underline{e}_d^i,n)\}. \tag{27}$$

$EP$ represents the expected returns from the production stage as seen by firm $i$ at the beginning of stage $d$.

We first note from (2) that

$$\Pr\{\tilde{A}_i(\underline{e}_d,n) = 1\} = \Pr\{\tilde{Q}_{dj}(e_{dj}) \leqslant \tilde{Q}_{di}(e_{di}) | \text{ for all } j = 1, \ldots, n\}, \tag{28}$$

or using the assumed independence of $\{\tilde{Q}_{dj} | j = 1, \ldots, n\}$

$$\Pr\{\tilde{A}_i(\underline{e}_d,n) = 1\} = \prod_{j=1} \Pr\{\tilde{Q}_{dj}(e_{dj}) \leqslant \tilde{Q}_{di}(e_{di})\}. \tag{29}$$

Thus, if $F_{dj}(q,e_{dj}) = \Pr\{\tilde{Q}_{dj}(e_{dj}) \leqslant q\}$ is the cumulative distribution function of $\tilde{Q}_{dj}(e_{dj})$, we may write (29) as

$$\Pr\{\tilde{A}_i(\underline{e}_d\, n) = 1\} = \prod_{j\neq i} F_{dj}(\tilde{Q}_{di}(e_{di}),e_{dj}). \tag{30}$$

*This striking result can be explained quite simply. If a firm is not concerned with follow-on benefits and is not directly rewarded for quality performance, it has no incentive to expend more than the minimum required effort during the production stage. This behavior cannot be altered by changing the method of competition for the production contract, but rather by (1) incorporating minimum acceptable levels of effort or quality into the production contract, (2) rewarding the firm directly for deliverying a quality system, and/or (3) encouraging the firm to believe that there are follow-on benefits of producing a quality system. See also the article by General Tashjian [16], who argues that in a two-stage (development and production) contract the government might motivate contractors to meet quality standards and the design-to-cost goal if it states in the development contract its intention to cancel the program if the design-to-cost goal is not met. As we point out in Section 6, determining the perceptions of contractors concerning future government actions (and hence, their follow-on benefits) is an important topic for future research.

Finally, using (30), (27) becomes

$$EP = \int_{-\infty}^{\infty} ([K_{pi} + h_{di}x] \prod_{j \neq i} F_{dj}(x, e_{dj})) f_{di}(x, e_{di})\, dx, \tag{31}$$

where $f_{di}(x, e_{di})$ is the probability density function of $\tilde{Q}_{di}(e_{di})$.

In the exponential case considered here, (31) becomes

$$EP = \frac{1}{q_{di}e_{di}} \int_0^{\infty} \left[ [K_{pi} + h_{di}x] \prod_{j \neq i} \left[ 1 - \exp\left( -\frac{x}{q_{dj}e_{dj}} \right) \right] \right] \exp\left( -\frac{x}{q_{di}e_{di}} \right) dx. \tag{32}$$

Restricting attention to $n = 1$ or 2, we obtain

$$EP(n=1) = \frac{1}{q_{di}e_{di}} \int_0^{\infty} [K_{pi} + h_{di}x] \exp\left( -\frac{x}{q_{di}e_{di}} \right) dx = K_{pi} + h_{di}q_{di}e_{pi}, \tag{33}$$

and setting $j \neq i$

$$\begin{aligned} EP(n=2) &= \frac{1}{q_{di}e_{di}} \int_0^{\infty} \left[ [K_{pi} + h_{di}x] \left[ 1 - \exp\left( 1 - \frac{x}{q_{dj}e_{dj}} \right) \right] \right] \exp\left( -\frac{x}{q_{di}e_{di}} \right) dx \\ &= [K_{pi} + h_{di}q_{di}e_{di}] \left( \frac{q_{di}e_{di}}{q_{dj}e_{dj} + q_{di}e_{di}} \right). \end{aligned} \tag{34}$$

Comparing (33) and (34), it is interesting to note that for any given level of effort during the development stage the ex ante expected returns from the production stage, which we denoted $EP$ above, are less for firm $i$ if 2 firms compete for the production contract than if firm $i$ alone is guaranteed the production contract.

Now, given (33)-(34), we may easily solve (26) for the optimal development effort $\hat{e}_{di}$, assuming the other firm's effort fixed at $e_{dj}$.

When $n = 1$, of course, there is no other competing firm and substituting (33) in (26) yields the following as the appropriate problem for firm $i$ (if firm $i$ is the only development firm):

$$\max_{e_{di} \geq 0} [G_d - e_{di}^2 + K_{pi} + h_{di}q_{di}e_{di}], \tag{35}$$

which has the unique solution

$$\hat{e}_{di} = \frac{h_{di}q_{di}}{2} \tag{36}$$

yielding overall profits for firm $i$ of

$$V_{di}(e_{di}, G_d) = G_d + K_{pi} + \frac{h_{di}^2 q_{di}^2}{4}. \tag{37}$$

When $n = 2$, matters are more complicated. Substitution of (34) in (26) yields

$$\max_{e_{di} \geq 0} \left[ [(G_d/2) - e_{di}^2 + [K_{pi} + h_{di}q_{di}e_{di}] \left( \frac{q_{di}e_{di}}{q_{dj}e_{dj} + q_{di}e_{di}} \right) \right]. \tag{38}$$

Taking first-order conditions in (38), while assuming $e_{dj}$ fixed, we obtain

$$e_{di} = \frac{K_{pi}q_{di}q_{dj}e_{dj}}{[2\Delta^2 - h_{di}q_{di}^2(\Delta + q_{dj}e_{dj})]} \tag{39}$$

where

$$\Delta = q_{d1}e_{d1} + q_{d2}e_{d2}. \tag{40}$$

We seek a Nash solution, defined by (8), which would be a simultaneous solution to (39) and the corresponding equation for firm $j$, i.e., to (39) and

$$e_{dj} = \frac{K_{pj}q_{dj}q_{di}e_{di}}{[2\Delta^2 - h_{dj}q_{dj}^2(\Delta + q_{di}e_{di})]}. \tag{41}$$

Assuming $\Delta$ fixed, and $h_{di}$, $h_{dj} > 0$, the simultaneous solution to (39) and (41) is

$$\hat{e}_{di}(\Delta) = \frac{-\Lambda}{[\Delta(2\Delta - h_{di}q_{di}^2)h_{dj}q_{di}q_{dj}^2 + h_{di}K_{pj}q_{di}^3q_{dj}^2]} \tag{42}$$

$$\hat{e}_{dj}(\Delta) = \frac{-\Lambda}{[\Delta(2\Delta - h_{dj}q_{dj}^2)h_{di}q_{di}^2q_{dj} + h_{dj}K_{pi}q_{di}^2q_{dj}^3]} \tag{43}$$

where

$$\Lambda = [K_{pi}K_{pj}q_{di}^2q_{dj}^2 - (2\Delta - h_{di}q_{di}^2)(2\Delta - h_{dj}q_{dj}^2)\Delta^2]. \tag{44}$$

Now, from (40) the Nash solution $\underline{\hat{e}}_d = \underline{\hat{e}}_d(\Delta)$ we seek must clearly satisfy (42)-(43) and

$$q_{d1}\hat{e}_{d1}(\hat{\Delta}) + q_{d2}\hat{e}_{d2}(\hat{\Delta}) = \hat{\Delta}. \tag{45}$$

Thus, multiplying (42) (respectively, (43)) by $q_{di}$ (respectively, $q_{dj}$) and adding the results leads to (45), which in general is a polynomial of degree 6 in the variable $\Delta$. Numerical solution procedures easily yield $\hat{\Delta}$ in general, and once $\hat{\Delta}$ is obtained so also is the desired Nash point $\underline{\hat{e}}_d$ from (41)-(42), from which all other desired information may be obtained. In this paper we will not proceed further with the general case. However, we discuss two cases which may be solved analytically.

CASE 1: $h_{di} = 0$ for all $i$. In this case (39)-(41) can be solved directly to yield

$$\hat{e}_{di} = T\sqrt{K_{pi}},\ \hat{e}_{dj} = T\sqrt{K_{pj}}, \tag{46}$$

where

$$T = \frac{(\sqrt{q_{d1}q_{d2}})(K_{p1}K_{p2})^{1/4}}{\sqrt{2}(q_{d1}\sqrt{K_{p1}} + q_{d2}\sqrt{K_{p2}})}. \tag{47}$$

In this case it can be shown that $\partial\hat{e}_{di}/\partial q_{di}$ has the same sign and $\partial\hat{e}_{di}/\partial K_{pj}$ has the opposite sign of $(q_{dj}\sqrt{K_{pj}} - q_{di}\sqrt{K_{pi}})$. As expected $\partial\hat{e}_{di}/\partial K_{pi} > 0$ always holds.

CASE 2: Two identical firms. When $q_{di} = q_{dj} = q_d$, $h_{di} = h_{dj} = h_d$, and $K_{pi} = K_{pj} = K_p$, we can again solve (39)-(41) explicitly, obtaining

$$\hat{e}_{d1} = \hat{e}_{d2} = \frac{3h_dq_d + \sqrt{9h_d^2q_d^2 + 32K_p}}{16}. \tag{48}$$

Here all the relative change effects are obvious and in the expected (positive) direction. An interesting point to note from (48) (or (46)-(47)) is that when $h_d = 0$, the amount of effort expended in development is independent of quality.

This concludes our discussion of Method 1 contracting (see (5)). Before considering further the government's problem in this regard, let us turn our attention briefly to Method 2 contracting (see (5')).

## 4. SOLUTION—METHOD 2

We continue to make the cost and distributional assumptions 1-5 of the previous section. In Method 2 contracting the stage $p$ behavior of the production contracting firm, say $i$, is determined as a solution to (5'), except that we further restrict $b$ so that $b \geqslant \underline{b} \geqslant 0$, with $\underline{b}$ being some minimal sharing rate set by Congress.* We assume the DoD utility function is specified linearly as

$$U_D(C,CO,Q) = -g_1 C - g_2 CO + g_3 Q, \tag{49}$$

where $g_i > 0$, $i = 1, 2, 3$. Then, for given $\alpha \in (0, 1)$, we may write the problem (5') as follows:

$$\begin{aligned} \underset{b,T_{pi},e_{pi}}{\text{Maximize}}\ EV &= \alpha E\{\tilde{\Pi}_{pi} + \tilde{F}_i\} + (1-\alpha) E\{U_D(\tilde{C}, \tilde{CO}, \tilde{Q}) | \tilde{Q}_{di} = Q_d\} \\ &= \alpha[(a+b)T_{pi} - be_{pi}^2 \\ &\quad + (H_i + h_{di}Q_d + h_{pi}q_{pi}e_{pi})] \\ &\quad + (1-\alpha)[-g_1(G_d + E\{\tilde{\Pi}_{pi} + \tilde{C}_{pi}(e_{pi})\}) \\ &\quad - g_2(G_d + E\{\tilde{\Pi}_{pi} + \tilde{C}_{pi}(e_{pi})\} - G) \\ &\quad + g_3(Q_d + q_{pi}e_{pi})]. \end{aligned} \tag{50}$$

Subject to: (4) and $\underline{b} \leqslant b \leqslant 1$.

Note that the expected total project cost (to the government) and quality (given $Q_d$) are, respectively, $G_d + E\{\tilde{\Pi}_{pi} + \tilde{c}_{pi}(e_{pi})\}$ and $Q_d + E\{\tilde{Q}_{pi}(e_{pi})\} = Q_d + q_{pi}e_{pi}$. Now we note that

$$E\{\tilde{\Pi}_{pi} + \tilde{C}_{pi}(e_{pi})\} = (a+b)T_{pi} + (1-b)e_{pi}^2. \tag{51}$$

Now, under our assumptions, (4) may be rewritten in the form (18). Moreover, as in Section 3, it may be shown here that for any fixed $b \in [\underline{b}, 1]$ the solution to (50) is on the boundary of the constraint set (18) provided only that†

$$\alpha > \frac{g_1 + g_2}{1 + g_1 + g_2}. \tag{52}$$

Condition (52) may be viewed as a lower bound on the bargaining power of firm $i$. We henceforth assume (52) so that (4) (i.e., (18)) holds as an equality. Just as in Section 3, we can now substitute (19) in (50) to obtain the final problem of interest:

$$\underset{b,T_{pi},e_{pi}}{\text{Maximize}}\ (-\alpha e_{pi}^2 + [\alpha h_{pi} + (1-\alpha)g_3]q_{pi}e_{pi} + Q_d[\alpha h_{di} + (1-\alpha)g_3] + TV(b)), \tag{53}$$

Subject to: $T_{pi} \geqslant 0$, $e_{pi} \geqslant 0$, $\underline{b} \leqslant b \leqslant 1$,

where the term $TV$ is independent of $e_{pi}$ and $Q_d$ and is given by

$$\begin{aligned} TV(b) &= \alpha H_i - (1-\alpha)g_1 G_d \\ &\quad + [\alpha - (1-\alpha)g_1]G_p \\ &\quad + [(1-\alpha)(g_1+g_2) - \alpha]k_{pi}(\gamma, b). \end{aligned} \tag{54}$$

*See also Canes [3], for a similar assumption and a discussion of some rationale for establishing such a lower bounding sharing rate.

†When (52) does not hold, the solution to (51) appears to be somewhat complicated as the solution need no longer be on the boundary of (18). Details for this more general case have not yet been worked out.

We may first note that (52) implies $[\alpha - (1-\alpha)(g_1 + g_2)] > 0$, and this coupled with (see (17)) $\partial k_{pi}/\partial b < 0$ implies that the optimal solution for $b$ in (53) is $\hat{b} = \underline{b}$ (note that the only term containing $b$ is $[\alpha - (1-\alpha)(g_1 + g_2)]\, k_p(\gamma, b)$. To obtain $e_{pi}$ we take first-order conditions in (53) and find

$$\hat{e}_{pi} = \text{Min}\left[\frac{[\alpha h_{pi} + (1-\alpha)g_3]q_{pi}}{2\alpha}\,,\ \left[\frac{G_p - k_p(\gamma,\underline{b})}{1-\underline{b}}\right]^{1/2}\right] \tag{55}$$

and $\hat{T}_{pi}$ is again found by substituting $\hat{e}_{pi}$ and $\hat{b} = \underline{b}$ into (19) to obtain

$$\hat{T}_{pi} = \frac{[G_p - k_p(\gamma,\underline{b}) - (1-b)\hat{e}_{pi}^2]}{(a+b)}. \tag{56}$$

Substituting $\hat{b} = \underline{b}$ and $\hat{e}_{pi}$ in (55) into (53), we see that Method 2 leads to exactly the same form of solution value (see 24) as Method 1 (where the $'$ denotes Method 2 values):

$$V'_{pi}(Q_d) = K'_{pi} + h'_{di}Q_d, \tag{57}$$

where for Method 2

$$K'_{pi} = -\alpha\hat{e}_{pi}^2 + [\alpha h_{pi}q_{pi} + (1-\alpha)g_3q_{pi}]\hat{e}_{pi} + TV(\underline{b}) \tag{58}$$

and

$$h'_{di} = [\alpha h_{di} + (1-\alpha)g_3]. \tag{59}$$

From this we see that the solution procedure and results for Method 1 in stage $d$ are completely transferable to Method 2, with $K'_{pi}$ and $h'_{di}$ substituted everywhere for $K_{pi}$ and $h_{di}$.

Before closing our analysis of Method 2 it is of interest to note, comparing (22) and (55), that effort expended in the production stage is always greater under Method 2 than under Method 1 contracting. More detailed comparative analysis of the other parameters and decisions will be explored in the next section via numerical analysis.

## 5. ILLUSTRATIVE RESULTS

We illustrate the concepts and results of the previous sections with a numerical example, solved in APL on the DEC System 10 at The Wharton School. We analyze the impact of the following parameters on the behavior of the firm and on the outcome of the project: variations in the risk sharing parameter (b) and partitioning of the (fixed) total government budget between the production budget ($G_p$) and the development budget. We consider three industry configurations: two identical firms, two firms with different levels of productivity, and a single monopolistic firm. The values of the parameters used in these experiments are given in Table 1. One can interpret these figures by assuming that money is measured in dollars and that quality is measured in miles of range of the weapon (e.g., an aircraft or missile). Simulations are run for Method 1 and for Method 2 with $\alpha = .8$ and $\alpha = .9$. A sample of the output for the two identical firms with Method 1 appears in Table 2. The remaining analyses are based on similar outputs for the other cases.

The impacts of the negotiation process (b) and budget allocated to production ($G_p$) on costs, quality, and the mean and variance of profit are shown in Table 3. These relationships are identical across all three industry structures. The mean and variance of cost to the government is the same for Method 1 and Method 2, whatever the value of $\alpha$. However, the expected quality and the expected cost to the firms are higher for Method 2 than for Method 1, and within Method 2 they are higher when $\alpha$ is at its lower value. In addition, as the expected cost of the firm increases from method 1 to Method 2, the expected profit (including intangibles) decreases. These effects occur, because Method 2 gives the

TABLE 1 — *Parameters in Experiment*

| Firm Parameters | | | |
|---|---|---|---|
| | Separate Firms | | Two Identical Firms |
| Parameters | Firm 1 | Firm 2 | and One Firm |
| $q_p$ | 1.6 | 1.5 | 1.55 |
| $h_p$ | 12,000 | 10,000 | 11,000 |
| $h_d$ | 1200 | 1000 | 1100 |
| $q_d$ | .8 | .6 | .7 |
| $\sigma$ | $10^7$ | $10^7$ | $10^7$ |
| $\eta$ | 1500 | 1500 | 1500 |
| $\delta$ | .7 | .7 | .7 |
| $\mu$ | $-10^6$ | $-10^6$ | $-10^6$ |
| Government Parameters | | | |
| $G = 1.2 \times 10^8$; $g_1 = g_2 = 1$; $g_3 = 10^4$; $\gamma = .15$; $a = .1$ | | | |

TABLE 2 — *Simulation Outputs for Method I with Two Identical Firms*

| Value of $b$ | Development Quality | Production Quality | Total Quality | Cost to Government | Standard Deviation of Cost to Government | Production Target |
|---|---|---|---|---|---|---|
| .1 | 4.242 | 11.627 | 15.869 | 110.64 | 9.00 | 0 |
| .3 | 4.295 | 13.214 | 17.509 | 112.72 | 7.00 | 4.6177 |
| .5 | 4.329 | 13.214 | 17.543 | 114.80 | 5.00 | 3.0770 |
| .7 | 4.363 | 13.214 | 17.577 | 116.88 | 3.00 | 4.3847 |
| .9 | 4.397 | 13.214 | 17.611 | 118.96 | 1.00 | 5.1692 |

| Value of $b$ | Cost to Firms | Profit of Firms | Standard Deviation of Profit of Firms | Intangible Profits | Development Effort | Production Effort |
|---|---|---|---|---|---|---|
| .1 | 88.92 | 151.73 | 16.1 | 130.01 | 4.0404 | 7.5011 |
| .3 | 106.14 | 154.08 | 14.9 | 147.50 | 4.0903 | 8.5250 |
| .5 | 106.68 | 155.65 | 13.8 | 147.53 | 4.1231 | 8.5250 |
| .7 | 107.21 | 157.22 | 13.0 | 147.55 | 4.1556 | 8.5250 |
| .9 | 107.75 | 158.78 | 12.5 | 147.58 | 4.1879 | 8.5250 |

Cost and profit are measured in millions of dollars, quality is measured in thousands of units, and effort is measured in thousands of units. $G_p = \$6 \times 10^7$ and $\gamma = 15\%$.

TABLE 3 — *Comparison of Methods for Any Industry Structure*

| Valid for any value of $b$ and $G_p$ | | | |
|---|---|---|---|
| Variable | Method 1 | Method 2 | |
| | | $\alpha = .9$ | $\alpha = .8$ |
| Expected Cost to the Government | Same | Same | Same |
| Expected Total Quality | Low | Medium | High |
| Expected Cost to the Firm | Low | Medium | High |
| Expected Profits of the Firm | High | Medium | Low |
| Intangible Reward | Low | Medium | High |
| Variance of Profit* | High | Medium | Low |
| Variance of Cost to the Government | Same | Same | Same |

*For the monopolistic industry structure, this variable is constant.

government more bargainning power than Method 1, and this bargainning power increases with decreasing $\alpha$. Thus, we obtain highest expected cost of firm and lowest expected profit at $\alpha = 0.8$.

The impact of the risk sharing parameter (b) and the portion of the budget allocated to production ($G_p$) on these variables are shown in Table 4. With regard to the risk sharing parameter, the results are what one would expect, with one exception. As $b$ increases, for any industry structure, the development and production efforts of each firm increases as long as the production target of the firm is zero. The production effort remains constant, with increasing $b$, once the target becomes positive. That is, as $b$ decreases, each firm attempts to respond by decreasing its target without changing its production effort. But as the target is constrained to be nonnegative, the firm meets the design-to-cost goal by decreasing its production effort. We also observe that when industry structure is monopolistic, development effort is independent of $b$. This occurs because the monopoly firm, assured of the contract, puts forth minimal effort at development stage to obtain intangible follow-on benefits. As a result, quality of the weapon, cost to the government, cost to the firm, profit of the firm (including intangible rewards), intangible rewards, and target increase (weakly) with increase in $b$. In addition, the variance of the government cost decreases with increasing $b$, since the firm assumes more risk. However, the variance of the firm's profit also decreases as the firm assumes increasing risk and this is an interesting result.

The explanation of this counterintuitive result arises partly from the fact that production cost and quality are correlated (which introduces a negative term in the variance calculation whose derivative may be dominant) and partly from the assumptions and parameter values used in these experiments. We begin by noting that the variance of profit (including intangibles) for the firm is given by

$$\mathrm{Var}\,(\tilde{\Pi}_i) = b^2\sigma_{pi} - 2bh_{pi}\sigma_{pi}\eta_{pi}\,\delta + h_{pi}^2\eta_{pi}^2 + h_{di}^2 q_{di}^2 e_{di}^2 + \mathrm{Var}(\tilde{C}_{di})$$

and thus,

$$\frac{d\,\mathrm{Var}\,(\tilde{\Pi}_i)}{db} = 2b\sigma_{pi}^2 - 2h_{pi}\sigma_{pi}\eta_{pi}\delta + 2h_{di}^2 q_{di}^2 e_{di}\frac{de_{di}}{db} + \frac{d\,\mathrm{Var}\,(\tilde{C}_{di})}{db}.$$

The variance of profit to the firm will increase or decrease with $b$ as the above result is positive or negative. For the parameter values used in these experiments, the result will always be negative as long as $\omega$, the coefficient of variation of $\tilde{C}_{di}$, is below .7 and will always be positive for $\omega > 2$. For intermediate values, the variance of $\tilde{\Pi}_i$ will decrease for low values of $b$ and will thereafter increase.

TABLE 4 — *Impact of $b$, $G_p$, and Industry Structure on Variables*

| Effects on variables as $b$ and $G_p$ change | | | | | | |
|---|---|---|---|---|---|---|
| | As $b$ Increases | | | As $G_p$ Increases | | |
| Variable | Different Firms | Identical Firms | One Firm | Different Firms | Identical Firms | One Firm |
| Expected Total Quality | INC | INC | LIM | INC | INC | LIM |
| Expected Cost to the Government | INC | INC | INC | SAME | SAME | SAME |
| Expected Cost to the Firms | INC | INC | LIM | INC | INC | INC |
| Expected Profit of the Firms | INC | INC | INC | CHG | CHG | LIM* |
| Expected Intangible Rewards | CHG | INC | LIM | CHG | INC | LIM |
| Production Target | INC | INC | INC | INC | INC | INC |
| Expected Quality at Development | INC | INC | SAME | INC | INC | SAME |
| Expected Quality at Production | CHG | LIM | LIM | CHG | LIM | LIM |
| Variance of Cost to the Government | DEC | DEC | DEC | SAME | SAME | SAME |
| Variance of Profit of the Firm | DEC | DEC | DEC | INC | INC | SAME |

*This applies only to Method 1. For Method 2, the profit increases and then decreases.

LEGEND:

INC: increases
DEC: decreases
SAME: no change
CHG: increases for small values, followed by decrease
LIM: increases until $T_p$ becomes positive, followed by no change

We also note that the variance of cost to the government is given by $(1 - b)^2 \sum_i \sigma_{pi}^2 p_i$, where $p_i$ is the probability that the $i^{\text{th}}$ firm will receive the production contract. When $\sigma_{pi}$ is independent of $i$ which is the case here, the variance of cost to the government is $(1 - b)^2 \sigma_p^2$, and this will always decrease with $b$.

Because increase in the risk sharing parameter bring about strict increases in total cost and total quality, risk sharing allows the government to obtain a cost/quality trade-off consistent with its goals. However, one may expect that beyond a certain point, quality increases slowly with $b$ (and in the case of the monopoly firm, will not increase at all), while cost continues to increase proportional to increasing $b$. This is the point where the target, or one or more of the targets for nonidentical firms, becomes positive.

The situation is more complex when the total government budget is partitioned between the development and production stages. We note that total quality increases as $G_p$ increases, because an

increase in $G_p$ relaxes the design-to-cost constraint. But, expected cost to government and the variance of that cost remain constant. Thus, the incentive to the government is to make $G_p$ as large as possible. However, there may be a limit to the size of $G_p$—that is, it may not be possible to let $G_p$ equal $G$—because the firms may refuse to compete due to inadequate compensation during the development stage.*

With regard to industry structure, in many cases variables such as cost, quality, and profit have the same relationship to each other for all values of $b$ and/or for all values of $G_p$. This is illustrated in Table 5. The government receives the highest quality weapon at the same or at lower cost when two different firms are competing for the contract, whereas it receives the lowest quality system at the same or at higher price when the industry structure is monopolistic. This is an expected result. Expected profit is highest for the monopoly firm and is lowest when the industry consists of two identical firms. The two-firm industry receives more profit when the firms are different, because one of the two firms is less productive than the other, and the decrease in expected profit for this firm is not offset by the increase in expected profit for the more productive firm. In general, one would expect the benefits of competition to diminish as the quality of the inefficient firm decreases. When the inefficient firm has low quality, the efficient firm will not perceive a credible competitive threat. So, increasing government's expenses on development does not repay in increased quality.

TABLE 5 — *Impact of Industry Structure on Variables for all Values of* $b$ and of $G_p$

| | For all values of $b$ | | | For all values of $G_p$ | | |
|---|---|---|---|---|---|---|
| Variable | Two Different Firms | Two Identical Firms | One Firm | Two Different Firms | Two Identical Firms | One Firm |
| Expected Quality | H | M | L | H | M | L |
| Expected Cost to the Government | S | S | H | S | S | S |
| Expected Cost to the Firms | H | M | L | F | F | F |
| Expected Profit of the Firms | M | L | H | M | L | H |
| Intangible Reward (Method 1) | H | L | M | H | M | L |
| Intangible Reward (Method 2) | H | M | L | F | F | L |
| Variance of Profit | H | M | L | H | M | L |
| Variance of Cost to the Government | S | S | S | S | S | S |

LEGEND:
H: high value
M: medium value
L: low value
S: same value
F: fluctuates

*We assume that firms have an incentive to participate in the development stage, even if their expected profit is negative or is below their alternative earnings opportunities. For example, the firms may wish to "stay in the game" and remain visible to DoD, a type of follow-on benefit not quantified in our model. However, within this constraint they wish to maximize expected profit plus follow-on benefits, which will lead the winner of the production contract to minimize his effort during the production stage when these follow-on benefits are nonexistent (i.e., when $h_{pi} = 0$). As we note in Section 6, examining the effects of removing this assumption is an important topic for future research.

The results described above are based on the assumptions that all firms in the industry will compete for the production contract regardless of the expected profits and that the government will allow all firms in the industry to compete (by paying a fixed cost for development) regardless of the productivity of the firm. These assumptions do not give rise to anomalies within the parameter ranges used here, contrary to expectation. For example, we observe that an increase in $G_p$ (and a corresponding decrease in the funds paid to the firms for their development efforts) results in an increase in quality with no change in the expected cost to the government, and also results in an increase in expected cost to the firms, regardless of industry structure. However, in reality, most firms have alternative uses for their resources (current and fixed assets, experienced managers, skilled workers, etc.), and some of them may decline to participate, once their expected profits do not compare favorably with those obtainable elsewhere. In fact, the existence of a negative constant term $K_{pi}(\gamma, b)$ in (23) can result in negative expected profits for suitable contract parameter values. This will almost certainly cause a firm to withdraw from participation at the development stage.

Generally, the selection of government contract parameters ($G_p$, $\gamma$, $b$, and $a$) must be compatible with the alternative earning opportunities of the firm. Such alternative market opportunities determine a set of contract parameters for each firm at which the firm would be willing to participate in the development effort and compete for the production contract. The government will desire a possibly different set of values for the contract parameters, for which the cost is low, variance of cost is low and the quality is high. The government must choose its values of $G_p$, $\gamma$, $b$, and $a$ from the set of contract parameter values which will guarantee prospective contractors earnings opportunities as attractive as their alternative market opportunities.

Figure 1 illustrates an example of this tradeoff. The figure shows a situation where the efficient and inefficient firms expect returns higher than \$102,300,000 and \$62,800,000, respectively. This expectation leads the firm to accept $b$ and $G_p/G$ only in specified ranges (as shown by the shaded area). The government has to choose its acceptable $b$ and $G_p/G$ depending on budget allocation, risk sharing, and other parameters. If these parameter values of the government fall in the shaded area, the government can expect the firms to bid for the project. If not, the government may have to revise its policy.

## 6. CONCLUSIONS AND FURTHER RESEARCH

We have examined the impact of dynamics and hierarchy, including industry structure, on government and industry behavior in a design-to-cost context, using a model of the weapons acquisition incentive process. Most of our results are a quantitative verification of what we would qualitatively expect. As risk sharing increases, the firms put forth more effort (except for the monopolistic firm during the development stage) and produce a better quality weapon at higher cost. The major counterintuitive result is that the variance of the government's cost decreases, and for some values of the economic and technological parameters, the variance of profit of the firm decreases as the risk sharing by the firm increases. In addition, the government receives a higher quality weapon for a fixed budget when it deals with competitive firms than it does when it must deal with a single monopolistic firm, and within limits, the quality increases when bidding firms are diverse in their capabilities. The government also receives a higher quality product when it invests a higher percentage of its budget in production (relative to development), unless the shift in investment causes some firms to withdraw from the competitive development phase of the acquisition process.

The interaction between government policy and industry structure suggests a productive direction for further research.* Government policies influence industry structure and also affect the structure of

*We are grateful to the referee for pointing out that DoD Directive 5000.28 on Design to Cost is undergoing revision. It is possible that the research reported here or future research of the type suggested in this section will be of use in formulating these and other government acquisition policies.

| | Case | Profit of Firm $ |
|---|---|---|
| Inefficient Firm | I1 | $0.59 \times 10^8$ |
| | I2 | $0.628 \times 10^8$ |
| | I3 | $0.65 \times 10^8$ |
| Efficient Firm | E1 | $10^8$ |
| | E2 | $1.023 \times 10^8$ |
| | E3 | $1.04 \times 10^8$ |

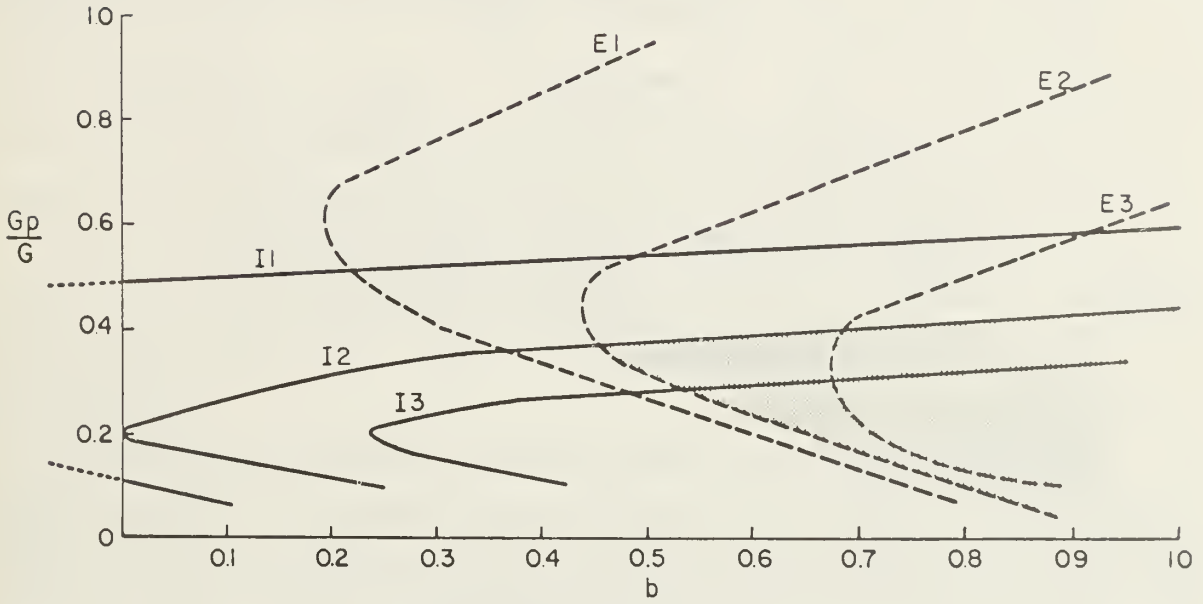



FIGURE 1. Areas of Firm Acceptability

the subset of the industry participating in the acquisition process and possibly in the long-term, the structure of the defense industry as a whole. In order to model the interaction between government policy and industry structure, we must determine three characteristics of the firms in the industry. The first characteristic, as mentioned in the previous section, is the spectrum of alternative earnings opportunities of the firms. The availability of these opportunities may lead some firms to withdraw from the development stage of the acquisition process. In contrast, the opportunities may lead other firms to participate in the hope that resources acquired during the project (skilled workers, experienced managers, etc.) may be useful in other areas, thus increasing the follow-on benefits of the project. The second characteristic is the range of risks that these opportunities present to the firms, which may lead some firms to decline to participate in the development phase of the project. The third is the firms' perceptions of future government actions (such as future projects, design-to-cost goals, and incentive parameters), and especially, of the uncertainty associated with these actions. It has been posited that a major blocking factor in industrial innovation is industry perception of the uncertainty in future government regulations and specifically that "the uncertainty of federal requirements, rather than their stringency, was perceived as the most important blocking factor,"* Thus, government parameter setting behavior, established during a sequence of projects, may induce perceptions and uncertainties about future government actions that will influence significantly the attractiveness of defense contracting to individual firms and thereby the structure of the defense industry.

*See Myers and Sweezy [13], page 29, for the quotation.

## ACKNOWLEDGMENT

The authors gratefully acknowledge the helpful comments of Marvin Denicoff of the Office of Naval Research (ONR) on an earlier draft of this paper. This research was carried out under Contract N00014-77-C-0171 of ONR.

## REFERENCES

[1] "Application of Design-to-Cost Concept to Major Weapon Systems Acquisitions," Report to the Congress by the Comptroller General of the United States, U.S. General Accounting Office, Washington, D.C., June 23, 1975.

[2] Bonin, J.P., "On the Design of Managerial Incentive Structures in a Decentralized Planning Environment," The American Economic Review, *66*, 4, 682-687 (September 1976).

[3] Canes, M.E., "The Simple Economics of Inventive Contracting: Note," The American Economic Review, *65*, 3, 478-483 (June 1975).

[4] Cummins, M.T., "Incentive Contracting for National Defense: A Problem of Optimal Risk Sharing," The Bell Journal of Economics, *8*, 1, 168-185 (Spring 1977).

[5] Defense Management Journal, "Design to Cost," Special Issue, *10*, 4, 1-40 (September 1974).

[6] Groves, T., "Incentives in Teams," Econometrica, *41* 4, 617-631 (1973).

[7] Groves, T., "Information Incentives and the Internalization of Production Externalities," in *Theory and Measurement of Economic Externalities*, S. Lin, Editor (The Academic Press, New York, 1976).

[8] Harris, M. and A. Raiv, "Some Results on Incentive Contracts with Applications to Education and Employment, Health Insurance, and Law Enforcement," The American Economic Review, *68*, 1, 20-30 (March 1978).

[9] Jennergren, P.L., "On the Design of Incentives in Business Firms—A Survey of Some Research," Management Science, *26* 2, 180-201 (1980).

[10] Kleindorfer, P.R. and M.R., Sertel, "Profit-Maximizing Design of Enterprises through Incentives," Journal of Economic Theory, *20*, 3, 318-339 (June 1979).

[11] McCall, J.J., "The Simple Economics of Incentive Contracting," The American Economic Review, *60*, 5, 837-846 (December 1970).

[12] McCullough, D.J., "Design to Cost Problem Definition, Survey of Potential Actors and Observations on Limitations," Paper P-928, Institute of Defense Analysis, Washington, D.C., January 1973.

[13] Myers, S. and E.E. Sweezy, "Why Innovations Falter and Fail: A Study of 200 Cases," Report R 75-04, Denver Research Institute, University of Denver, January, 1976.

[14] Office of Management and Budget. Circular A-109, April, 1976.

[15] Ross, S.A., "On the Economic Theory of Agency and the Principle of Similarity," in *Essays on Economic Behavior under Uncertainty* M. Balch, D. McFadden, and S. Wu, Editors (North Holland, Amsterdam, 1974).

[16] Tashjian, M.T., "Implementation of the Design-to-Cost Concept from the Contractual Point of View," Defense Management Journal, *10*, 4, 8-17 (September 1974).

# OPTIMAL STRATEGIES IN A GAME OF ECONOMIC SURVIVAL

Karl Borch

*Norwegian School of Business*
*Bergen-Sandviken, Norway*

**ABSTRACT**

The games of economic survival introduced by Shubik and Thompson seem tailor-made for the analysis of some problems in insurance and have found many applications in this industry. The optimal strategy in such games may be a so-called "band strategy." This result seems counter-intuitive and has caused some puzzlement. This paper gives sufficient conditions so that the optimal strategy will be of a simpler form, and it is argued that these conditions are satisfied in most applications to insurance.

## INTRODUCTION

The term "economic survival game" is really another name for the classic problem of the gambler's ruin which can be traced back to Pascal. The new term is usually associated with Shubik and Thompson [11], who extended some of the classic results to games of strategy. This generalization in itself may not have been particularly fruitful, but it led to an increased interest in the multi-period one-person game against nature. It is the starting point of the papers by Miyasawa [7] and Morrill [9], and the general ideas are behind the book by Dubins and Savage [4]. A few years earlier De Finetti [3] presented a similar model, as a generalization of the older actuarial theories, based on the probability that an insurance company shall become insolvent or "ruined."

Miyasawa and Morrill both show that the optimal strategy can be a "band strategy"—a concept to be explained in the next section. Strategies of this form are not intuitively appealing and do not seem to agree well with observations when the model is applied to situations in real life, but they have been studied for their mathematical interest by several authors including Hallin [6]. The purpose of this note is to show that a band strategy can be optimal only in fairly exceptional cases.

The presentation is given in terms of insurance, which was the application Morrill [8] originally had in mind. This interpretation was also the starting point of De Finetti, and his models have been developed and discussed extensively in actuarial literature, by, among others, Borch [1], Bühlmann [2] and Pentikäinen [10]. It is, of course, possible to give other interpretations to the model.

## THE PROBLEM

De Finetti's starting point is an insurance company which underwrites identical portfolios in each consecutive operating period. Let the portfolio be characterized by:

$P =$ premiums received in an operating period.

$x =$ claims paid during the period, a stochastic variable with the distribution $F(x)$.

Let $S_t$ be the company's equity capital or "surplus" at the end of period $t$. Clearly, $S_t$ follows a random walk defined by:

$$S_{t+1} = S_t + P - x.$$

In insurance it is natural to assume that if $S_t < 0$ the company is ruined and is not allowed to operate in the following periods.

At the end of period $t$, the company may consider paying a dividend, $s_t \in (0, S_t)$. The problem is then to find the optimal dividend policy. De Finetti studied different assumptions about the company's objectives. One of these was that the company would seek to maximize the expected discounted sum of the dividend payments:

$$E\left\{\sum_{t=0}^{\infty} \mathrm{v}^t s_t\right\} \tag{1}$$

where $\mathrm{v} \in (0.1)$ is the discount factor.

Let $V(S)$ denote the maximum, i.e., the expected discounted sum of the dividend payments if the initial equity capital is $S$ and if the company follows an optimal dividend policy. It is easy to see that $V(S)$ must satisfy the functional equation:

$$V(S) = \max_{0 \leqslant s \leqslant S} \left\{ s + \mathrm{v} \int_0^{S+P-s} V(S - s + P - x)\, dF(x) \right\}. \tag{2}$$

Essentially, early dividend payments are preferred, but a high dividend implies a low retained surplus and hence a high probability of ruin, i.e., that the payments will terminate. One would therefore expect that an optimal $s$ exists for any value of $S$.

Equation (2) can be solved by standard methods of dynamic programming, and Miyasawa and Morrill have shown that the optimal dividend policy is of the form:

$$\begin{aligned}
&s = 0 && \text{for} && S \leqslant Z_1 \\
&s = S - Z_1 && \text{for} && Z_1 < S \leqslant Z_2 \\
&s = 0 && \text{for} && Z_2 < S \leqslant Z_3 \\
&\ldots && \ldots && \ldots \\
&s = S - Z_{2n+1} && \text{for} && Z_{2n+1} < S.
\end{aligned}$$

A policy of this form defined by a set of $2n + 1$ numbers, is called a "band strategy" by Morrill. The special case with $n = 0$ he calls a "barrier strategy."

A band strategy implies that a company may omit or reduce its dividend after a successful operating period, and this seems to be contradicted by the business behavior one can observe. The barrier strategy does not invite the same objections. Essentially, it implies that the company decides that it is optimal to maintain a certain equity capital $Z$. No dividend is paid if the equity capital is below this level. If the equity capital exceeds $Z$, the excess is paid out as dividend immediately. This seems reasonable and suggests that, in normal cases, a barrier strategy may be optimal.

Assume now that the company decides to adopt a barrier strategy defined by a number $Z$, not necessarily optimal. The expected discounted value of the dividend payments under this policy is

$$V(S, Z) = \mathrm{v} \int_0^{S+P} V(S + P - x, Z)\, dF(x). \tag{3}$$

If a density, $f(x) = F'(x)$, exists, this is an integral equation with the boundary conditions

$$V(S, Z) = 0 \text{ for } S < 0 \text{ and } V(S, Z) = S - Z + V(Z, Z) \text{ for } S \geqslant Z.$$

The equation has a unique continuous solution for any nonnegative value of $Z$. Morrill's result implies that the solutions in $Z$ to the equation

$$\frac{\partial V(S, Z)}{\partial Z} = 0 \tag{4}$$

are independent of $S$. A barrier strategy will be optimal only if (4) has a unique root in $Z$.

## A SOLUTION OF THE PROBLEM

The integral equation (3) can be solved, and it is possible to study the roots of (4) by this direct approach. It seems, however, more convenient to take a roundabout way and consider a rather naive dividend policy. Assume that at the end of a profitable operating period the company pays the whole profit out as dividend immediately. If the period has given a loss, no dividend is paid.

Let $W(S)$ be the expected discounted sum of the dividends paid under this policy when the company's initial capital is $S$: The function $W(S)$ is clearly a solution of the integral equation:

$$W(S) = \int_0^P (P-x)\, f(x)\, dx + \mathrm{v} W(S) \int_0^P f(x)\, dx + \mathrm{v} \int_P^{S+P} W(S+P-x)\, f(x)\, dx \tag{5}$$

which evidently is far simpler than the functional equation (2). Setting the first term on the right-hand side in (5) equal to $C$, and rearranging the other terms, the equation takes the form

$$\{1 - \mathrm{v}F(P)\} W(S) = C + \mathrm{v} \int_0^S W(S-x)\, f(x+P)\, dx. \tag{6}$$

Taking the Laplace transform of (6) one finds:

$$\{1 - \mathrm{v}F(P)\} \int_0^\infty W(S) e^{-tS}\, dS = \frac{C}{t} + \left\{ \mathrm{v} \int_0^\infty W(x) e^{-tx}\, dx \right\} \left\{ \int_0^\infty f(x+P) e^{-tx} dx \right\}.$$

Writing

$$1 - \mathrm{v}F(P) = K$$

and

$$\int_0^\infty f(x+P) e^{-tx}\, dx = \phi(t)$$

the equation takes the form

$$t \int_0^\infty W(x) e^{-tx}\, dx = \frac{C}{K - \mathrm{v}\phi(t)}. \tag{7}$$

The left-hand side of (7) can be written

$$t \int_0^\infty W(x) e^{-tx}\, dx = W(0) + \int_0^\infty W'(x) e^{-tx}\, dx.$$

As $\mathrm{v}\phi(t) \leqslant \phi(0) = 1 - F(P) < 1 - \mathrm{v}F(P) = K$ for $\mathrm{v} < 1$, it follows that the right-hand side of (7) can be expanded in a convergent series. Hence:

$$\int_0^\infty W'(x)e^{-\alpha x}dx + W(0) = \frac{C}{K}\sum_{n=0}^{\infty}\left(\frac{v}{K}\phi(t)\right)^n. \tag{8}$$

By taking $S = 0$ in (6) one sees that $W(0) = \frac{C}{K}$, so that (8) can be written

$$\int_0^\infty W'(x)e^{-\alpha x}dx = W(0)\sum_{n=1}^{\infty}\left(\frac{v}{K}\phi(t)\right)^n. \tag{9}$$

$(\phi(t))^n$ is the Laplace transform of the $n$th convolution of $f(x+p)$ with itself, which will be denoted by $f_n(x)$. Hence, $f_n(x) = \int_0^x f_{n-1}(y)f(x+P-y)dy$ for $n > 1$, and let by definition $f_1(x) = f(x+P)$.

Taking the inverse transform of (9), one obtains:

$$W'(x) = W(0)\sum_{n=1}^{\infty}\left(\frac{v}{K}\right)^n f_n(x). \tag{10}$$

From (10) $W(x)$ itself can be found by integration:

$$W(x) = W(0)\sum_{n=0}^{\infty}\left(\frac{v}{K}\right)^n F_n(x).$$

This result can be verified directly. To show this, introduce a distribution $G(x)$ defined as follows:

$$G(0) = F(P)$$
$$G(x) = F(x+P) \quad \text{for } x > 0.$$

Substitution of this in the calculations above gives

$$W(x) = C\{1 + vG(x) + v^2 G_2(x) \ldots\}.$$

$G_n(x)$ is the probability that accumulated losses in $n$ periods shall not exceed the initial equity capital $x$. If this event occurs, the expected discounted value of the next dividend payment is $v^n C$.

The convolutions $G_n(x)$ have a complicated form, and this symbol will not be used in the following.

The expansion (10) is, as noted, convergent, and $W(x)$ as a discounted sum of distribution functions is a bounded and nondecreasing function of $x$. From (6) it follows that

$$W(0) = \frac{C}{1 - vF(P)}.$$

Further, it is easy to see without any calculation, that

$$W(\infty) = \frac{C}{1-v} = \sum_{t=0}^{\infty} v^t \int_0^P (P-x)f(x)dx.$$

The equation says that if the equity capital is infinite, the company will never be ruined, and the expected dividend will be equal to expected profit in all future periods.

From (10) it follows that

$$W'(0) = W(0)\frac{\nu}{K} f(P)$$

and

$$W'(\infty) = 0.$$

## A SUFFICIENT CONDITION THAT A BARRIER STRATEGY IS OPTIMAL

Consider now the equation

$$W'(Z) = 1. \tag{11}$$

It can be shown that the solutions of (4) and (11) coincide. This is a tedious procedure, and one can arrive at the result by an indirect approach.

If $Z$ is a solution of (11) and if the company's capital at the end of an operating period is $Z + z$ with $z > 0$, we will have, at least for $z$, sufficiently small

$$W(Z + z) < W(Z) + z.$$

Hence, the expected discounted sum of the dividends paid will increase if the company departs from its naive policy and immediately pays the amount $z$ out as dividend.

Similarly, if the capital is smaller than $Z$, it will be optimal to retain profits until the capital has been brought up to $Z$ before paying any dividend.

A sufficient condition that equation (11) has a unique root is $W'(0) > 1$ and $W''(x) < 0$.

Differentiation of (10) gives

$$W''(x) = W(0) \sum_{n=1}^{\infty} \left(\frac{\nu}{K}\right)^n f_n'(x). \tag{12}$$

From the definition of the convolution, it follows that:

$$f_n(x) = \int_0^x f_{n-1}(y) f(x + P - y)\, dy \quad \text{for } n \geq 2.$$

Differentiation of this expression gives:

$$f_n'(x) = f(P) f_{n-1}(x) + \int_0^x f_{n-1}(y) f'(x + P - y)\, dy$$

and by definition

$$f_1'(x) = f'(x + P).$$

Substitution of these expressions into (12) gives

$$W''(x) = W(0)\frac{\nu}{K}\left\{f'(x+P) + \frac{\nu}{K} f(P) f(x+P)\right\} + W(0) \sum_{n=2}^{\infty} \left(\frac{\nu}{K}\right)^n \left\{\int_0^x f_{n-1}(y) f'(x + P - y)\, dy + \frac{\nu}{K} f(P) f_n(x)\right\}. \tag{13}$$

The first term in (13) is negative if

$$f'(x+P) + \frac{v}{K} f(P) f(x+P) < 0. \tag{14}$$

Substitution of

$$f'(x+P-y) = -\frac{v}{K} f(P) f(x+P-y)$$

in the following terms, shows that they all vanish. Hence, (14) is a sufficient condition that $W''(x) < 0$.

Condition (14) can be written:

$$\frac{f'(x+P)}{f(x+P)} < -\frac{v}{K} f(P)$$

and this gives

$$f(x+P) < f(P) e^{-\frac{v}{K} f(P) x}.$$

Hence, equation (11) will have a unique root if the claim density goes to zero at least as rapidly as an exponential.

A necessary, but not sufficient condition is that $f'(x+P) < 0$ for $x > 0$. For a unimodal distribution this condition will be satisfied when $P$ is greater than the mode. The typical claim distribution in insurance will be skew, with a long tail. For such distributions the mean will be greater than the mode, and the premium $P$ will again normally be greater than the mean, i.e., than expected claim payments.

If the necessary condition is satisfied, the sufficient condition (14) appears as a mild regularity condition for the tail of the distribution, and the curious cases found by Miyasawa and Morrill may be dismissed as interesting but not very relevant.

Clearly, the company will not find it optimal to maintain positive reserves unless

$$W'(0) > 1.$$

Substitution of the expressions derived from (10) gives the condition in the form:

$$vf(P) \int_0^P (P-x) f(x)\, dx > (1 - vF(P))^2.$$

Since the left-hand side will go to zero with $P$, it follows that the premium must be above a certain level to induce the company to risk any part of its equity capital as reserves in the underwriting business. The company will, however, not refuse to underwrite any portfolio, since it cannot lose unless it put some of its own money at risk.

## SPECIAL CASE

As a simple special case assume that the claim density is exponential, i.e.,

$$f(x) = e^{-x}.$$

Substitution in the general formula gives:

$$W(x) = \frac{C}{1-v}\left\{1 - \frac{ve^{-P}}{1-v+ve^{-P}} \exp\left[-\frac{(1-v)x}{1-v+ve^{-P}}\right]\right\}$$

and

$$W'(x) = \frac{Cve^{-P}}{(1-v+ve^{-P})^2} \exp\left[-\frac{(1-v)x}{1-v+ve^{-P}}\right].$$

It is easy to see that $W''(x) < 0$, and hence that a barrier strategy is optimal.

Equation (11) which determines the optimal strategy takes the form

$$\frac{1-v}{1-v+ve^{-P}} Z = \ln C + \ln v - P - 2\ln(1-v+ve^{-P}),$$

or by substituting the expression for $C$ and writing $v = (1+i)^{-1}$,

$$\frac{iZ}{i+e^{-P}} = \ln(P-1+e^{-P}) - 2\ln(1+i) - P - 2\ln(i+e^{-P}).$$

The table below gives the value of $Z$ for some selected values of $P$ when $i = 0.1$.

| Premium $P$ | Optical Reserve $= Z$ | Ruin probability $e^{-Z-P}$ |
|---|---|---|
| 1.0 | Negative | — |
| 1.2 | Negative | — |
| 1.3 | 0.166 | 0.23 |
| 1.4 | 1.052 | 0.086 |
| 1.5 | 1.42 | 0.054 |
| 2.0 | 2.28 | 0.014 |
| 3.0 | 2.14 | 0.0059 |
| 4.0 | 1.47 | 0.0042 |
| 7.0 | 0 | 0.0009 |

The table shows that increasing $P$ will increase the quality of the insurance contracts, i.e., reduce the probability of ruin. It should not be surprising that $Z$ will decrease with increasing $P$ for large values of $P$. If the level of premiums is very high, there will be a very low probability that the underwriting shall lead to a loss, and the incentive to maintain additional reserves will be reduced.

## RELEVANCE TO INSURANCE

The barrier strategy may be less objectionable than the band strategy, but it still does not agree very well with the observations one can make of dividend policies in real life. If an insurance company has had an exceptionally profitable period, it will usually pay out only a part of the profits as dividend. Sometimes this is justified by an explicit statement to the effect that it is desirable to retain some profits in order to safeguard future dividend payments. This clearly contradicts the assumption that the company's objective is to maximize (1).

Even if the company and its shareholders seem to agree that the objective should be to maximize an expression of the form (1), one may observe apparent departures from the barrier strategy. This may occur just because the model is too simple to give an adequate representation of the complex real world, but a number of other possible explanations suggest themselves, among other things:

(i) It may be considered fair that the insured, i.e., the policyholders, should receive some benefits after a very profitable period. One way of arranging this would be to increase the company's reserves, and thus reduce the probability of ruin in the next period, i.e., improve the security of the policy holders. In some countries the governments seem to try to induce insurance companies to take this attitude and maintain a conservative dividend policy.

(ii) The managers of an insurance company may rightly or wrongly believe that they will be blamed for sharp reductions in dividend payments. Their natural reaction may then be to argue in favor of a conservative dividend policy which, incidentally, will increase the expected life of the company and, hence, the job security of managers and other employees.

One will get a smoother sequence of dividend payments by assuming that the company's objective is to maximize

$$\sum_{t=0}^{\infty} v^t u(s_t) \tag{15}$$

where $u(s)$ is a concave utility function.

This generalization of De Finetti's model has been studied by several authors, by, among others, Hakansson [5], who finds that, for a particular class of utility functions, the optimal strategy is of the form $s = qS$. This means that at the end of each operating period the company pays out a fraction, $q$, of its equity capital as dividend, and represents behavior in conformance with observations. The objections to criterion function (15) are essentially of a theoretical nature. If the company should reduce the scale of its operations, or stop selling insurance altogether, it would find it optimal to pay out its equity capital in an infinite decreasing sequence. This might make sense in the consumption plans discussed by Hakansson but not for a dividend policy.

The discount factor v which appears in the formula, should be based on an interest rate which represents a pure risk premium. The company's reserves can be invested to earn a return, which could be paid directly to the shareholders as owners of the capital. Theoretically, they could even keep the capital, provided that they accept liability for possible losses in the insurance operations. Government regulations will usually require that the reserves should be kept in low-risk assets, which also must be of fairly liquid nature, since the reserves can be called upon at short notice to pay claims. This means that the rate of return on the reserves will be modest, probably close to the risk-free market rate of interest. The shareholders risk their capital as reserves in the insurance company and will require a higher expected return. This interest differential seems to be the one to use for determining the discount factor.

## REFERENCES

[1] Borch, K., *The Mathematical Theory of Insurance* (D.C. Heath & Co., Lexington, MA, 1974).
[2] Bühlmann, H., *Mathematical Methods of Risk Theory* (Springer Verlag, Berlin, 1970).
[3] De Finetti, B., "Su una Impostazione Alternativa della Teoria Collettiva del Rischio," Transactions of the XV International Congress of Actuaries, *2*, 433-443 (1957).
[4] Dubins, L.E. and L.J. Savage, *How to Gamble if You Must* (McGraw-Hill, New York, 1965).
[5] Hakansson, N., "Optimal Investment and Consumption Strategies under Risk for a Class of Utility Functions," Econometrica, 587-607 (1970).
[6] Hallin, M., "Band Strategies: The Random Walk of Reserves," Blätter der DGVM, *14*, 231-236 (1979).
[7] Miyasawa, K., "An Economic Survival Game," Journal of the Operations Research Society of Japan, 95-113 (1962).
[8] Morrill, J., "Discrete Economic Survival Game Models for Insurance Surplus Distribution," Unpublished dissertation, University of Michigan, Ann Arbor, MI, 1964.

[9] Morrill, J., "One-Person Games of Economic Survival," Naval Research Logistics Quarterly, 49-69 (1966).
[10] Pentikäinen, T., "A Model of Stochastic-Dynamic Prognosis: An Application of Risk Theory to Business Planning." Scandinavian Actuarial Journal, 29-53 (1975).
[11] Shubik, M. and G.L. Thompson, "Games of Economic Survival," Naval Research Logistics Quarterly, 111-123 (1959).

# DIAGNOSTIC ANALYSIS OF INVENTORY SYSTEMS: AN OPTIMIZATION APPROACH*

Gabriel R. Bitran, Arnoldo C. Hax and Josep Valor-Sabatier

*Sloan School of Management*
*Massachusetts Institute of Technology*
*Cambridge, Massachusetts*

**ABSTRACT**

The objective of a diagnostic analysis is to provide a measure of performance of an existing system and estimate the benefits of implementing a new one, if necessary. Firms expect diagnostic studies to be done promptly and inexpensively. Consequently, collection and manipulation of large quantities of data are prohibitive. In this paper we explore aggregate optimization models as tools for diagnostic analysis of inventory systems. We concentrate on the dynamic lot size problem with a family of items sharing the same setup, and on the management of perishable items. We provide upper and lower bounds on the total cost to be expected from the implementation of appropriate systems. However, the major thrust of the paper is to illustrate an approach to analyze inventory systems that could be expanded to cover a wide variety of applications. A fundamental by-product of the proposed diagnostic methodology is to identify the characteristics that items should share to be aggregated into a single family.

## 1. INTRODUCTION

Inventory-control theory provides a wide variety of models to manage effectively products with different characteristics. Extensive surveys are available that culminate in a taxonomy identifying specific decision rules to manage inventories under a number of conditions. (Silver [7], Nahmias [5], and Aggarwal [1].) A very legitimate concern of many managers today is to understand the extent to which these models could contribute to the enhancement of the performance of their current inventory system. Consequently, management-science practitioners are frequently faced with requests to estimate the benefits for improving the performance of such systems. Firms expect that the estimation process in itself will not represent a major project and consume a considerable amount of human and financial resources; the analysis must be reasonably accurate and inexpensive.

Two basic approaches have emerged to comply with those requirements. One, the statistical approach, is based on analyzing the inventory performance on a relatively small sample of items and generating inferences that can be applied to the whole population. The other approach, based on optimization techniques, consists in devising simple models to estimate the benefits of implementing the proposed system.

The simplified models are derived by aggregating items into families, thereby reducing the data collection and computation efforts. The ultimate objective is the generation of bounds relating the aggregated model to the original problem. Zipkin [12] and Evans [3] have developed such bounds for linear and transportation models, respectively.

*This research was performed under Contract N00014-75-C-0556 for the Office of Naval Research

The management-science literature has rarely addressed the identification of problems and the a-priori measurement of benefits to be derived from the implementation of a proposed new system. Recently Hax, Majluf and Pendrock [4], and Wagner [9] have stressed the importance of developing diagnostic analysis tools. In particular Hax, Majluf and Pendrock [4] report the results of a diagnostic analysis of a large logistics system.

In [2] we illustrated how statistically-based techniques, like clustering, inference and sampling, may be used in diagnostic studies. In this paper we focus on the application of aggregate-optimization models. In Section 2 we concentrate on the dynamic-lot-size problem with a family of items sharing the same setup. We examine an aggregation scheme and compute upper and lower bounds on the total cost to be expected from the implementation of appropriate systems. We also provide conditions under which the solution to the aggregate problem solves the original problem. In Section three we analyze the management of perishable items. An aggregate version of the newsboy problem is studied and bounds are determined. Conclusions and topics for future research are discussed in the last section.

We would like to emphasize that although this paper examines two classes of inventory models, our aim is to propose a general approach to design diagnostic methodologies that could be easily extendable to a wide spectrum of optimization models. Moreover, we hope that the result of our analysis not only contributes to the development of quantitative measures to evaluate inventory performance, but also provides insights into the conditions that have to be met by the items to become candidates for aggregation into a family unit. This last result could contribute to the important subject of aggregate planning.

## 2. LOT SIZE PROBLEMS: A DIAGNOSTIC ANALYSIS

In this section we develop diagnostic procedures to estimate ordering and inventory carrying costs associated with the operation and management of a family of items replenished in lots. The objective of this procedure is to obtain estimates of total costs while reducing as much as possible the massive manipulation of data and the computational requirements. We derive error bounds for aggregate problems and rules indicating how to aggregate the items in a family to minimize such errors.

We address practical situations in which all items in a family are replenished simultaneously and share the same ordering costs.

Whenever the items in the family satisfy the conditions underlying the economic order quantity system [6], the total annual cost of ordering and holding the family in stock is easily shown to be

$$TC = \sqrt{2Sr}\ \sqrt{\sum_{i=1}^{N} (D_i v_i)} \tag{2.1}$$

where $S$, $r$, $D_i$ and $v_i$ are, respectively, the ordering costs, the inventory carrying factor, the annual demand of item $i$, and the unit cost of item $i$. The cost of holding in stock one unit of product $i$ for one year is $r v_i$. Therefore, (2.1) can be rewritten as

$$TC = (\text{production cost of family})^{1/2}\ \sqrt{2Sr}.$$

In this case total ordering and holding cost for the family of items can be readily computed from aggregate data. However, when the assumptions underlying the economic order quantity system differ significantly from the actual conditions, the estimation of the costs must be based on more complex models.

In what follows we assume that the demand for each item in the family, in each period $1, 2, \ldots T$, is deterministic. The inventory control problem then, is to solve the following mixed integer programming problem:

$$(P) \qquad z_P = \min f(x) = \sum_{t=1}^{T} \left[ S_t \delta\left( \sum_{i=1}^{N} x_{i,t} \right) + \sum_{i=1}^{N} p_{i,t} x_{i,t} + \sum_{i=1}^{N} h_{i,t} \sum_{\tau=1}^{t} (x_{i,\tau} - d_{i,\tau}) \right]$$

subject to

$$\sum_{\tau=1}^{t} (x_{i,\tau} - d_{i,\tau}) \geqslant 0 \quad t = 1, 2, \ldots, T; \; i = 1, 2, \ldots, N.$$

$$x_{i,\tau} \geqslant 0 \quad \tau = 1, 2, \ldots, T; \; i = 1, 2, \ldots, N.$$

$$\delta\left( \sum_{i=1}^{N} x_{i,t} \right) = \begin{cases} 1, & \text{if } \sum_{i=1}^{N} x_{i,t} > 0 \\ 0, & \text{if } \sum_{i=1}^{N} x_{i,t} \leqslant 0 \end{cases}$$

$$t = 1, 2, \ldots, T; \; i = 1, 2, \ldots, N.$$

where $p_{i,t}$, $h_{i,t}$, $d_{i,t}$, $x_{i,t}$ and $S_t$ are, respectively, the unit cost, holding cost, demand and amount ordered for item $i$ in period $t$, and the ordering cost in period $t$. Without loss of generality, we assume that the initial inventory is zero.

We will compare problem (P) with the solution obtained from the following aggregate problem:

$$(AP) \qquad z_{AP} = \min \varphi(y) = \sum_{t=1}^{T} \left[ S_t \delta(y_t) + p_t y_t + h_t \sum_{\tau=1}^{t} (y_\tau - d_\tau) \right]$$

subject to

$$\sum_{\tau=1}^{t} (y_\tau - d_\tau) \geqslant 0 \quad t = 1, 2, \ldots, T$$

$$y_\tau \geqslant 0 \quad \tau = 1, 2, \ldots, T$$

$$\delta(y_t) = \begin{cases} 1, & \text{if } y_t > 0 \\ 0, & \text{if } y_t \leqslant 0 \end{cases} \quad t = 1, 2, \ldots, T$$

where $p_t$, $h_t$, $d_t$, $y_t$ and $S_t$ are the unit cost, holding cost, the demand and amount ordered for the family in period $t$, and the ordering cost in period $t$. We assume that $y_t$ and $d_t$ are defined as follows for conveniently selected constants $k_i > 0$, $i = 1, 2, \ldots, N$:

$$(2.2) \qquad y_t = \sum_{i=1}^{N} k_i x_{i,t} \quad t = 1, 2, \ldots, T, \; i = 1, 2, \ldots, N.$$

$$(2.3) \qquad d_t = \sum_{i=1}^{N} k_i d_{i,t} \quad t = 1, 2, \ldots, T, \; i = 1, 2, \ldots, N.$$

Therefore, if $x_{i,t}$ $i = 1, 2, \ldots, N$; $t = 1, 2, \ldots, T$ is feasible in (P), $y_t$ $t = 1, 2, \ldots, T$ defined as in (2.2), is feasible in $(AP)$. Wagner and Whitin [10] provide an efficient dynamic programming algorithm to solve $(P)$ and $(AP)$.

The expression below, for a pair $(x, y)$ satisfying (2.2) and (2.3), is useful in our development.

$$f(x) - \varphi(y) = \sum_{t=1}^{T} S_t \left[ \delta\left( \sum_{i=1}^{N} x_{i,t} \right) - \delta(y_t) \right] + \sum_{t=1}^{T} \left[ \sum_{i=1}^{N} h_{i,t} \sum_{\tau=1}^{t} (x_{i,\tau} - d_{i,\tau}) - h_t \sum_{\tau=1}^{t} (y_\tau - d_\tau) + \sum_{i=1}^{N} p_{i,t} x_{i,t} - p_t y_t \right]$$

$$= \sum_{t=1}^{T} \sum_{i=1}^{N} \sum_{\tau=1}^{t} (h_{i,t} - h_t k_i)(x_{i,\tau} - d_{i,\tau}) + \sum_{t=1}^{T} \sum_{i=1}^{N} x_{i,t}(p_{i,t} - k_i p_t). \tag{2.4}$$

The second equality follows from (2.2), (2.3) and the fact that

$$\delta\left( \sum_{i=1}^{N} x_{i,t} \right) = \delta(y_t) \quad t = 1, 2, \ldots, T.$$

PROPOSITION 2.1 (Theorem 6 in [8]): $(P)$ and $(AP)$ have at least one optimal solution of the form

$$\sum_{\tau=1}^{t} (x_{i,\tau} - d_{i,\tau})\, x_{i,t+1} = 0 \quad t = 1, 2, \ldots, T;\ i = 1, 2, \ldots, N \tag{2.5}$$

$$\sum_{\tau=1}^{t} (y_\tau - d_\tau) y_{t+1} = 0 \quad t = 1, 2, \ldots, T. \tag{2.6}$$

Note that (2.5) and (2.6) characterize the extreme points in $(P)$ and $(AP)$, respectively. The proposition indicates that in order to compare $z_P$ and $z_{AP}$, it is sufficient to consider feasible solutions of $(P)$ and $(AP)$ such that in each period, either the initial inventory is zero, and eventually an order is placed, or the initial inventory is positive and no order is placed.

PROPOSITION 2.2: Let $y_t$ $t = 1, 2, \ldots, T$ be feasible in $(AP)$ and assume it satisfies (2.6). Then, there is a unique set $x_{i,t}$ $t = 1, 2, \ldots, T$; $i = 1, 2, \ldots, N$ feasible in $(P)$ such that $y_t$ and $x_{i,t}$ are related as in (2.2) and $x_{i,t}$ satisfies (2.5).

Whenever $x_{i,t}$ and $y_t$ $i = 1, 2, \ldots, N$, $t = 1, 2, \ldots, T$ are as in Proposition 2.2 we will say that they are corresponding feasible extreme point solutions.

Unfortunately, if $y_t$ $t = 1, 2, \ldots, T$ is an optimal extreme point in $(AP)$ the corresponding feasible extreme point solution in $(P)$ need not be optimal, in general, in that problem. In fact, it is not even true, in general, that the optimal value of one of the problems is a lower or upper bound of the other. Two instances where a relation between $z_P$ and $z_{AP}$ is known *a priori* are given in the next two results.

PROPOSITION 2.3: If $k_i, p_t$ and $h_t$ are chosen such that $p_{i,t} - p_t k_i \leqslant 0$, $h_{i,t} - h_t k_i \leqslant 0$ $(p_{i,t} - p_t k_i \geqslant 0,\ h_{i,t} - h_t k_i \geqslant 0)$ for all $i$ and all $t$, it follows that

$$z_{AP} \geqslant z_P \ (z_{AP} \leqslant z_P).$$

PROOF: Follows by Proposition 2.2 and Expression (2.4). □

COROLLARY 2.1: Assume the unit and holding costs $p_{i,t}$, $h_{i,t}$ $i = 1, 2, \ldots, N$ and $t = 1, 2, \ldots, T$ remain in the same proportion for all $t$, i.e., $p_{i,t} = k_i p_t$, $h_{i,t} = k_i h_t$ for some $k_i > 0$, $p_t > 0$ and $h_t > 0$. Then, $z_P = z_{AP}$ as long as $p_t$ and $h_t$ are used in $(AP)$ as the underlying unit and holding costs.

PROOF: Follows by Proposition 2.3. □.

A special case of Corollary 2.1, of interest in practical settings, is when the unit and holding costs are constant for each item over the planning horizon and the holding cost is proportional to the unit cost. In these situations $p_{i,t} = p_i, h_{i,t} = h_i$ and $h_i = rp_i$ for all items and therefore $z_P = z_{AP}$.

Even when the conditions of corollary 2.1 do not hold, we may prefer to solve $(AP)$ instead of $(P)$ either to save computational work or because family demand data is readily available or even because family demand can be forecast with greater accuracy than item demand. $k_i$ and $h_t$ can always be chosen to satisfy the conditions in Proposition 2.3. However, the gap between $z_P$ and $z_{AP}$ can be greater than desired. In what follows we attempt to measure this aggregation gap under quite general conditions.

In the remainder of this section we assume the following:

$$k_i = 1,\ D_i = \sum_{t=1}^{T} d_{i,t},\ D = \sum_{i=1}^{N} D_i,\ c_i = \frac{D_i}{D},$$

$$D_t = \sum_{i=1}^{N} d_{i,t},\ d_{i,t} = c_i D_t + \varphi_{i,t},\ p_t = \sum_{i=1}^{N} c_i p_{i,t} \text{ and } h_t = \sum_{i=1}^{N} c_i h_{i,t}$$

where, for all expressions

$$i = 1, 2, \ldots, N;\ t = 1, 2, \ldots, T.$$

Note that

$$\sum_{i=1}^{N} \varphi_{i,t} = \sum_{t=1}^{T} \varphi_{i,t} = 0 \tag{2.7}$$

for all $i$ and $t$.

Let $x$ be feasible in $(P)$ and let

$$\mathcal{R}(x) = \left\{ t : 1 \leqslant t \leqslant T,\ \sum_{i=1}^{N} x_{i,t} > 0 \right\}.$$

For $t \in \mathcal{R}(x)$ we denote by $n(x, t)$ the smallest integer in $\mathcal{R}(x) \cup \{T + 1\}$ larger than $t$, i.e.,

$$n(x, t) = \min\{\tau : \tau > t,\ \tau \in \mathcal{R}(x) \cup \{t + 1\}\}.$$

$\mathcal{R}(x)$ and $n(x, t)$ correspond, respectively, to the set of time periods when any production occurs and the next such period after $t$.

Assume $\bar{x}$ is an optimal extreme point of $(P)$. Then,

$$z_P = \sum_{t=1}^{T} S_t \delta\left(\sum_{i=1}^{N} \bar{x}_{i,t}\right) + \sum_{t \in \mathcal{R}(\bar{x})} \sum_{i=1}^{N} \left[ p_{i,t} \sum_{\tau=t}^{n(\bar{x},t)-1} d_{i,\tau} + \sum_{\tau=t+1}^{n(\bar{x},t)-1} d_{i,\tau} \sum_{u=t}^{\tau-1} h_{i,u} \right].$$

Let $\bar{y}$ be the corresponding feasible extreme point solution to $\bar{x}$ in $(AP)$ then,

$$\begin{aligned} z_P &= \sum_{t=1}^{T} S_t \delta(\bar{y}_t) + \sum_{t \in \mathcal{R}(\bar{x})} \sum_{i=1}^{N} \left[ \sum_{\tau=t}^{n(\bar{x},t)-1} p_{i,t} d_{i,\tau} + \sum_{\tau=t+1}^{n(\bar{x},t)-1} d_{i,\tau} \sum_{u=t}^{\tau-1} h_{i,u} \right] \\ &= \sum_{t=1}^{T} S_t \delta(\bar{y}_t) + \sum_{t \in \mathcal{R}(\bar{x})} \sum_{i=1}^{N} \left[ \sum_{\tau=t}^{n(\bar{x},t)-1} p_{i,t}(c_i D_\tau + \varphi_{i,\tau}) + \sum_{\tau=t+1}^{n(\bar{x},t)-1} (c_i D_\tau + \varphi_{i,\tau}) \sum_{u=t}^{\tau-1} h_{i,u} \right] \end{aligned}$$

$$= \sum_{t=1}^{T} S_t \delta(\bar{y}_t) + \sum_{t \in \mathscr{R}(\bar{x})} \left[ p_t D_t + \sum_{\tau=t+1}^{n(\bar{x},t)-1} D_\tau \left( p_t + \sum_{u=t}^{\tau-1} h_u \right) \right]$$

$$+ \sum_{t \in \mathscr{R}(\bar{x})} \sum_{i=1}^{N} \left[ p_{i,t} \varphi_{i,t} + \sum_{\tau=t+1}^{n(\bar{x},t)-1} \varphi_{i,\tau} \left( p_{i,t} + \sum_{u=t}^{\tau-1} h_{i,u} \right) \right] \geqslant z_{AP} + \Phi(\bar{x})$$

where

$$\Phi(\bar{x}) = \sum_{t \in \mathscr{R}(\bar{x})} \sum_{i=1}^{N} \left[ p_{i,t} \varphi_{i,t} + \sum_{\tau=t+1}^{n(\bar{x},t)-1} \varphi_{i,\tau} \left( p_{i,t} + \sum_{u=t}^{\tau-1} h_{i,u} \right) \right]. \tag{2.8}$$

We have just proved

PROPOSITION 2.4: If $\bar{x}$ is an optimal extreme point of $(P)$,

$$z_P \geqslant z_{AP} + \Phi(\bar{x}).$$

Similarly, define for any extreme point $y$ in $(AP)$

$$\mathscr{R}(y) = \{t : 1 \leqslant t \leqslant T,\ y_t > 0\}$$

and

$$n(y, t) = \min\{\tau : \tau > t,\ t \in \mathscr{R}(y) \cup \{T+1\}\}.$$

Let $\bar{y}$ be an optimal extreme point of $(AP)$ and $\tilde{x}$ be its corresponding feasible extreme point solution in $(P)$. Then,

$$z_{AP} = \sum_{t=1}^{T} S_t \delta(\bar{y}_t) + \sum_{t \in \mathscr{R}(\bar{y})} \left[ p_t D_t + \sum_{\tau=t+1}^{n(\bar{y},t)-1} D_\tau \left( p_t + \sum_{u=t}^{\tau-1} h_u \right) \right] = \sum_{t=1}^{T} S_t \delta \left( \sum_{i=1}^{N} \tilde{x}_{i,t} \right)$$

$$+ \sum_{t \in \mathscr{R}(\bar{y})} \sum_{i=1}^{N} \left[ p_{i,t}(d_{i,t} - \varphi_{i,t}) + \sum_{\tau=t+1}^{n(\bar{y},t)-1} \left( p_{i,t} + \sum_{u=t}^{\tau-1} h_{i,u} \right) (d_{i,\tau} - \varphi_{i,\tau}) \right] \geqslant z_P - \Gamma(\bar{y})$$

where

$$\Gamma(\bar{y}) = \sum_{t \in \mathscr{R}(\bar{y})} \sum_{i=1}^{N} \left[ p_{i,t} \varphi_{i,t} + \sum_{\tau=t+1}^{n(\bar{y},t)-1} \varphi_{i,\tau} \left( p_{i,t} + \sum_{u=t}^{\tau-1} h_{i,u} \right) \right]. \tag{2.9}$$

Consequently,

PROPOSITION 2.5: If $\bar{y}$ is an optimal extreme point of $(AP)$,

$$z_P \leqslant z_{AP} + \Gamma(\bar{y}).$$

COROLLARY 2.2: If $\bar{y}$ and $\bar{x}$ are optimal extreme point solutions of $(AP)$ and $(P)$, respectively, then

$$\Phi(\bar{x}) \leqslant z_P - z_{AP} \leqslant \Gamma(\bar{y}).$$

Note that $\bar{x}$ and $\bar{y}$ need not be corresponding feasible solutions in Corollary 2.2.

Corollary 2.1 shows that if the unit and holding costs, for each item in the family, remain in the same proportion, then $z_P = z_{AP}$. A similar result holds for the demands even when $h_t \neq \sum_{i=1}^{N} c_i h_{i,t}$ and/or $p_t \neq \sum_{i=1}^{N} c_i p_{i,t}$.

COROLLARY 2.3: Assume that $\varphi_{i,t} = 0$ $i = 1, 2, \ldots, N$, $t = 1, 2, \ldots, T$, that is, $d_{i,t} = c_i D_t$. Then

$$z_P = z_{AP}.$$

PROOF: From (2.8) and (2.9) we have $\Phi(\bar{x}) = \Gamma(\bar{y}) = 0$ and by Corollary 2.1, $0 \leqslant z_P - z_{AP} \leqslant 0$. Hence, $z_P = z_{AP}$. □

Corollary 2.3 provides an interesting interpretation of problem $(AP)$. More specifically, it shows that solving the aggregate problem is equivalent to solving the detailed problem with the demands $d_{i,t}$ replaced by $c_i D_t = \frac{D_i}{D} D_t$. In fact, under the conditions of the corollary, the following stronger result holds:

PROPOSITION 2.6: Under the conditions of Corollary 2.3, $f(x) = \varphi(y)$ for every pair $x, y$ of corresponding extreme points.

PROOF: Note that $\mathcal{R}(x) \equiv \mathcal{R}(y)$ and $n(x, t) = n(y, t)$ for all $t \in \mathcal{R}(x)$ and $\mathcal{R}(y)$.

$$\begin{aligned} f(x) &= \sum_{t=1}^{T} S_t \delta\left(\sum_{i=1}^{N} x_{i,t}\right) + \sum_{i=1}^{N} \sum_{t \in \mathcal{R}(x)} \left[p_{i,t} d_{i,t} + \sum_{\tau=t+1}^{n(x,t)-1} \left(\sum_{u=t}^{\tau-1} h_{i,u} + p_{i,t}\right) d_{i,\tau}\right] \\ &= \sum_{t=1}^{T} S_t \delta(y_t) + \sum_{i=1}^{N} \sum_{t \in \mathcal{R}(y)} \left[p_{i,t} c_i D_t + \sum_{\tau=t+1}^{n(y,t)-1} \left(\sum_{u=t}^{\tau-1} h_{i,u} + p_{i,t}\right) c_i D_\tau\right] \\ &= \sum_{t=1}^{T} S_t \delta(y_t) + \sum_{t \in \mathcal{R}(y)} \left[p_t D_t + \sum_{\tau=t+1}^{n(y,t)-1} \left(\sum_{u=t}^{\tau-1} h_u + p_t\right) D_\tau\right] = \varphi(y). \end{aligned}$$ □

We should note here that the conditions of Corollary 2.3 (proportionality of the demands) are equivalent to requiring the multiplicative seasonal factors of the different items to be equal period by period. In what follows we prove this result.

PROPOSITION 2.7: Assume that we fit the model $d_{i,t} = s_{i,t}(b_{0,i} + b_{1,i}t)$ to every item $i$, where $b_{0,i}$ and $b_{1,i}$ are determined by the least squares method, and the $s_{i,t}$ are the multiplicative seasonal factors defined according to Winters [11]. Then, if the demands are proportional as required in Corollary 2.3, it follows that $s_{i,t} = s_{l,t}$ for all $i = 1, 2, \ldots, N$, $l = 1, 2, \ldots, L$ and $t = 1, 2, \ldots, T$.

PROOF: For each of the $N$ items the offset value $b_{0,i}$ and the linear trend $b_{1,i}$ have the property of minimizing the sum of squares of the seasonal variation:

$$b_{1,i} = \frac{\sum_{t=1}^{T} \left(d_{i,t} - \frac{D_i}{T}\right)(t - \bar{t})}{\sum_{t=1}^{T} (t - \bar{t})^2} = A \sum_{t=1}^{T} (T d_{i,t} - D_i)(t - \bar{t})$$

$$b_{0,i} = \frac{D_i}{T} - b_{1,i}\bar{t}$$

where

$$A = \frac{\sum_{t=1}^{T} (t - \bar{t})^2}{T} \text{ and } \bar{t} = \frac{1 + T}{2}$$

the seasonal factors are

$$s_{i,t} = \frac{d_{i,t}}{b_{0,i} + tb_{1,i}}.$$

Assuming proportional demands: $d_{i,t} = k_{i,l} d_{l,t}$ $i = 1, 2, \ldots, N$ $t = 1, 2, \ldots, T$ for a generic item $i$ and a fixed "base" item $l$, we get

$$d_{i,t} = k_{i,l}(b_{0,l} + tb_{1,t})s_{l,t} = s_{l,t}(k_{i,l} b_{0,l} + tk_{i,l} b_{1,l}$$

but

$$b_{1,i} = A \sum_{t=1}^{T} (Td_{i,t} - D_i)(t - \bar{t}) = A \sum_{t=1}^{T} (Tk_{i,l} d_{l,t} - k_{i,l} D_l)(t - \bar{t})$$

$$= k_{i,l} A \sum_{t=1}^{T} (Td_{l,t} - D_l)(t - \bar{t}) = k_{i,l} b_{1,l}$$

and

$$b_{0,i} = \frac{D_i}{T} - b_{1,i}\bar{t} = \frac{k_{i,l} D_l}{T} - k_{i,l} b_{1,l}\bar{t} = k_{i,l} b_{0,l}.$$

Therefore,

$$d_{i,t} = s_{l,t}(b_{0,i} + tb_{1,i})$$

which implies:

$$s_{i,t} = s_{l,t} \quad i = 1, 2, \ldots, N \quad t = 1, 2, \ldots, T.$$ □

When the conditions of Corollary 2.3 do not hold it is possible to derive bounds for the gap between $z_P$ and $z_{AP}$ as we show in the following development.

Let $a_i$ be defined as the maximum absolute value of the relative difference between each period demand of product $i$ and its annual demand, and the corresponding ratio of the family, i.e.,

$$a_i = \max_{\substack{t=1,2,\ldots,T \\ t: d_{i,t}>0}} \frac{\left| \frac{d_{i,t}}{D_t} - \frac{D_t}{D} \right|}{\frac{d_{i,t}}{D_t}} \quad i = 1, 2, \ldots, N. \tag{2.10}$$

Hence,

$$a_i = \max_{\substack{t=1,2,\ldots,T \\ t: d_{i,t}>0}} \left| 1 - \frac{c_i D_t}{d_{i,t}} \right| = \max_{\substack{t=1,2,\ldots,T \\ t: d_{i,t}>0}} \left| 1 - \frac{d_{i,t} - \varphi_{i,t}}{d_{i,t}} \right| \tag{2.11}$$

$$= \max_{\substack{t=1,2,\ldots,T \\ t: d_{i,t}>0}} \frac{|\varphi_{i,t}|}{d_{i,t}}.$$

Let

$$a = \max_{i=1,2,\ldots,N} a_i = \max_{i=1,2,\ldots,N} \left\{ \frac{|\varphi_{i,t}|}{d_{i,t}} : t = 1, 2, \ldots, T, \quad t : d_{i,t} > 0, \ i = 1, 2, \ldots, N \right\}.$$

Thus,

$$|\varphi_{i,t}| \leqslant a d_{i,t} \quad i = 1, 2, \ldots, N, \ t = 1, 2, \ldots, T \text{ and } d_{i,t} > 0.$$

Note that when $d_{i,t} = 0$, $\varphi_{i,t} = -c_i D_t \leqslant 0$.

For any extreme point $x$ in $(P)$,

$$|\Phi(x)| = \left| \sum_{t \in \mathcal{R}(x)} \sum_{i=1}^{N} \left[ \sum_{\tau=t}^{n(x,t)-1} p_{i,t} \varphi_{i,\tau} + \sum_{\tau=t+1}^{n(x,t)-1} \sum_{u=t}^{\tau-1} h_{i,u} \varphi_{i,\tau} \right] \right| \tag{2.12}$$

$$\leqslant \left| \sum_{t \in \mathcal{R}(x)} \sum_{i=1}^{N} \left[ \sum_{\substack{\tau=t \\ \tau : d_{i,\tau} > 0}}^{n(x,t)-1} p_{i,t} \varphi_{i,\tau} + \sum_{\substack{\tau=t+1 \\ \tau : d_{i,\tau} > 0}}^{n(x,t)-1} \sum_{u=t}^{\tau-1} h_{i,u} \varphi_{i,\tau} \right] \right|$$

$$\leqslant \sum_{t \in \mathcal{R}(x)} \sum_{i=1}^{N} \left[ \sum_{\substack{\tau=t \\ \tau : d_{i,\tau} > 0}}^{n(x,t)-1} p_{i,t} |\varphi_{i,\tau}| + \sum_{\substack{\tau=t+1 \\ \tau : d_{i,\tau} > 0}}^{n(x,t)-1} \sum_{u=t}^{\tau-1} h_{i,u} |\varphi_{i,\tau}| \right]$$

$$\leqslant a \sum_{t \in \mathcal{R}(x)} \sum_{i=1}^{N} \left[ \sum_{\tau=t}^{n(x,t)-1} p_{i,t} d_{i,\tau} + \sum_{\tau=t+1}^{n(x,t)-1} \sum_{u=t}^{\tau-1} h_{i,u} d_{i,\tau} \right] = a [P_P(x) + H_P(x)]$$

where $P_P(x)$ and $H_P(x)$ are the proportional purchasing and holding costs in problem $(P)$ associated with the feasible point $x$.

Similarly, for any extreme point $y$ in $(AP)$

$$|\Gamma(y)| \leqslant a \sum_{t \in \mathcal{R}(y)} \sum_{i=1}^{N} \left[ \sum_{\tau=t}^{n(y,t)-1} p_{i,t} d_{i,\tau} + \sum_{\tau=t+1}^{n(y,t)-1} \sum_{u=t}^{\tau-1} h_{i,u} d_{i,\tau} \right]. \tag{2.13}$$

However,

$$d_{i,\tau} - c_i D_\tau \leqslant |d_{i,\tau} - c_i D_\tau| = |\varphi_{i,\tau}| \leqslant a d_{i,\tau}.$$

Thus, if $a < 1$,

$$d_{i,\tau} \leqslant \frac{1}{1-a} c_i D_\tau. \tag{2.14}$$

Substituting (2.14) in (2.13)

$$|\Gamma(y)| \leqslant \frac{a}{1-a} \sum_{t \in \mathcal{R}(y)} \left[ \sum_{\tau=t}^{n(y,t)-1} \sum_{i=1}^{N} p_{i,t} c_i D_\tau + \sum_{i=1}^{N} \sum_{\tau=t+1}^{n(y,t)-1} \sum_{u=t}^{\tau-1} h_{i,u} c_i D_\tau \right] \tag{2.15}$$

$$= \frac{a}{1-a} [P_{AP}(y) + H_{AP}(y)]$$

where $P_{AP}(y)$ and $H_{AP}(y)$ are the proportional costs in $(AP)$ corresponding to $y$.

Before deriving the desired bounds we observe that if $\bar{x}$ and $\bar{y}$ denote, respectively, optimal extreme points of $(P)$ and $(AP)$ then $z_P$ and $z_{AP}$ can be written as

$$z_P = S_P(\bar{x}) + P_P(\bar{x}) + H_P(\bar{x})$$

and

$$z_{AP} = S_{AP}(\bar{y}) + P_{AP}(\bar{y}) + H_{AP}(\bar{y})$$

where $S_P(\bar{x})$ and $S_{AP}(\bar{y})$ are the fixed costs associated with $\bar{x}$ and $\bar{y}$. Also, to use as reference value in the measurement of the gap, we define

$$z^x_{\max} = \sum_{t=1}^{T}\left[S_t + \sum_{i=1}^{N} p_{i,t}d_{i,t}\right] + \max\{P_P(x) + H_P(x): x \text{ feasible in } (P)\} = S^x + (P^x + H^x)_{\max}$$

and

$$z^y_{\max} = \sum_{t=1}^{T}(S_t + p_tD_t) + \max\{P_{AP}(y) + H_{AP}(y): y \text{ feasible in } (AP)\} = S^y + (P^y + H^y)_{\max}.$$

Therefore, by Proposition 2.5 and Expression (2.15), if $a < 1$

$$\frac{z_P - z_{AP}}{z^y_{\max} - z_{AP}} \leqslant \frac{|\Gamma(\bar{y})|}{S^y - S_{AP}(\bar{y}) + (P^y + H^y)_{\max} - P_{AP}(\bar{y}) - H_{AP}(\bar{y})}$$

$$\leqslant \frac{a}{1-a}\,\frac{H_{AP}(\bar{y}) + P_{AP}(\bar{y})}{S^y - S_{AP}(\bar{y}) + (P^y + H^y)_{\max} - P_{AP}(\bar{y}) - H_{AP}(\bar{y})}.$$

Since $(P^y + H^y)_{\max} \geqslant P_{AP}(\bar{y}) + H_{AP}(\bar{y}) \geqslant 0$ and

$$S^y \geqslant z_{AP} = S_{AP}(\bar{y}) + P_{AP}(\bar{y}) + H_{AP}(\bar{y})$$

it follows that

$$\frac{z_P - z_{AP}}{z^y_{\max} - z_{AP}} \leqslant \frac{a}{1-a}. \tag{2.16}$$

Expression (2.16) provides an upper bound on the gap between the optimal values $z_P$ and $z_{AP}$ relative to $z^y_{\max}$ provided that $a < 1$. Another gap measurement often employed in this type of analysis is obtained considering only the aggregate problem as a basis for comparison, i.e.,

$$\frac{z_P - z_{AP}}{z_{AP}} \leqslant \frac{|\Gamma(\bar{y})|}{z_{AP}} \leqslant \frac{\frac{a}{1-a}[P_{AP}(\bar{y}) + H_{AP}(\bar{y})]}{S_{AP}(\bar{y}) + P_{AP}(\bar{y}) + H_{AP}(\bar{y})} \leqslant \frac{a}{1-a} \tag{2.17}$$

Other upper bounds can be obtained as follows:

$$\frac{z_{AP} - z_P}{z^x_{\max} - z_{AP}} \leqslant \frac{-\Phi(\bar{x})}{S^x + (P^x + H^x)_{\max} - S_P(\bar{x}) - P_P(\bar{x}) - H_P(\bar{x})} \leqslant \frac{a[P_P(\bar{x}) + H_P(\bar{x})]}{S^x - S_P(\bar{x})} \leqslant a. \tag{2.18}$$

The inequalities in (2.18) are implied by Proposition 2.4, Expression (2.12) and the facts that $(P^x + H^x)_{\max} \geqslant P_P(\bar{x}) + H_P(\bar{x}) \geqslant 0$ and $S^x - S_P(\bar{x}) \geqslant P_P(\bar{x}) + H_P(\bar{x})$.

Similarly,

$$\frac{z_{AP} - z_P}{z_P} \leqslant \frac{-\Phi(\bar{x})}{S_P(\bar{x}) + P_P(\bar{x}) + H_P(\bar{x})} \leqslant a. \tag{2.19}$$

Lower bounds on the value of $z_P$ are readily obtainable from inequalities (2.18) and (2.19). These expressions imply (2.20) and (2.21) below:

$$z_P \geqslant \frac{z_{AP} - a z^x_{\max}}{1 - a} \text{ if } a < 1 \tag{2.20}$$

and

$$z_P \geqslant \frac{z_{AP}}{1 + a}. \tag{2.21}$$

Note that if the production costs $p_{i,t}$ do not change over time, they can be dropped from the objective function in the formulation of problems $(P)$ and $(AP)$. In addition, the reference costs $z^x_{\max}$ and $z^y_{\max}$ can be redefined as

$$z^x_{\max} = \sum_{t=1}^{T} S_t + \max\{H_P(x) : x \text{ feasible in } (P)\}$$

$$z^y_{\max} = \sum_{t=1}^{T} S_t + \max\{H_{AP}(y) : y \text{ feasible in } (AP)\}$$

in order to get tighter bounds.

We summarize the results obtained in the next proposition.

PROPOSITION 2.8: The relative gap between $z_P$ and $z_{AP}$ can be bounded as follows:

$$(1)\quad \frac{z_P - z_{AP}}{z^y_{\max} - z_{AP}} \leqslant \frac{a}{1 - a} \text{ if } a < 1 \tag{2.16}$$

$$(2)\quad \frac{z_P - z_{AP}}{z_{AP}} \leqslant \frac{a}{1 - a} \text{ if } a < 1 \tag{2.17}$$

$$(3)\quad \frac{z_{AP} - z_P}{z^x_{\max} - z_P} \leqslant a \tag{2.18}$$

$$(4)\quad \frac{z_{AP} - z_P}{z_P} \leqslant a \tag{2.19}$$

$$(5)\quad \frac{z_P - z_{AP}}{z_{AP}} \geqslant \frac{a}{1 - a}\frac{z_{AP} - z^x_{\max}}{z_{AP}} \text{ if } a < 1 \tag{2.22}$$

$$(6)\quad \frac{z_P - z_{AP}}{z_{AP}} \geqslant -\frac{a}{1 + a} \tag{2.23}$$

Expressions (2.22) and (2.23) are equivalent to (2.20) and (2.21), respectively.

Proposition 2.8 indicates that $(AP)$ is a good approximation to $(P)$ whenever the value of $a$ is "small." When all data on the demands of the items are known, the value of $a$ can be computed exactly. In this case if the bounds are satisfactory we may want to solve $(AP)$ to reduce the computational requirements. However, most likely, in practical settings the $d_{i,t}$'s are not known exactly and $a$ can not be calculated. Proposition 2.8 can still be very useful in these occasions because the value of $a$ depends only on the maximum relative absolute deviation of the demands of the items and of the family, and managers are often able to estimate bounds for such a value. An important by-product of Proposition 2.8 is the qualitative insight it provides for aggregating items in diagnostic studies of the nature discussed in this paper.

Finally, we point out that (2.18) and (2.22) hold if $z^z_{\max}$ is replaced by an upper bound. The same is true with (2.16) with respect to $z^y_{\max}$. Also, if it is assumed that the demand of item $i$ has no trend, $a_i$ can be seen as the maximum absolute value of the relative difference between the seasonal factors of the item and the seasonal factors of the family.

## 3. SINGLE PERIOD STOCHASTIC DEMAND

In this section we address the following problem: Given a family of $N$ items with single period stochastic demands, determine upper and lower bounds on the maximum deviation between replenishing the items individually and as an aggregate family.

The following conditions below are assumed to hold throughout this section:

(i) The interval of time with nonzero demand is the same for all items.

(ii) The demand of each item is independent of the demand of the other items and is normally distributed with mean $\mu_i$ and standard deviation $\sigma_i$.

(iii) The unit under stock and over stock costs are proportional to the value $v_i$ of the item and are written as

over-stock unit cost $\equiv co_i = c_1 v_i \quad i = 1, 2, \ldots, N.$

under-stock unit cost $\equiv cu_i = c_2 v_i \quad i = 1, 2, \ldots, N.$

In order to determine the desired bounds it is necessary to establish the following results.

PROPOSITION 3.1: Consider the problem

$$\hat{z} = \max \sum_{i=1}^{N} \sigma_i$$

subject to

$$\sum_{i=1}^{N} \sigma_i^2 = \sigma^2$$

$$\sigma_i \geqslant 0 \qquad i = 1, 2, \ldots, N.$$

Then

$$\hat{z} = \sqrt{N}\sigma.$$

PROPOSITION 3.2:

Consider the problem

$$(W) \qquad \bar{z} = \min \sum_{i=1}^{N} \sigma_i$$

subject to

$$\sum_{i=1}^{N} \sigma_i^2 = \sigma^2$$

$$\underline{\sigma} \leqslant \sigma_i \leqslant \bar{\sigma} \qquad i = 1, 2, \ldots, N$$

where we assume that $\underline{\sigma} < \bar{\sigma}$. When $\underline{\sigma} = \bar{\sigma}$ the resolution of $(W)$ is trivial. $\sigma^2$, $\underline{\sigma}$ and $\bar{\sigma}$ are known nonnegative numbers.

Then,

$$\bar{z} = N_1\underline{\sigma} + N_2\bar{\sigma} + \gamma\,(\sqrt{\sigma^2 - N_1\underline{\sigma}^2 - N_2\bar{\sigma}^2}),$$

where

$$N_1 = \left\lceil \frac{N\bar{\sigma}^2 - \sigma^2}{\bar{\sigma}^2 - \underline{\sigma}^2} \right\rceil,$$

$$\gamma = \left\lfloor \frac{\sigma^2 - N_1\underline{\sigma}^2}{\bar{\sigma}^2} - \left\lceil \frac{\sigma^2 - N_1\underline{\sigma}^2}{\bar{\sigma}^2} \right\rceil \right\rfloor$$

and

$$N_2 = N - N_1 - \gamma.$$

$\lceil x \rceil$ and $\lfloor x \rfloor$ denote, respectively, the largest integer less or equal than $x$ and the smallest integer greater or equal than $x$.

PROOF: Problem $(W)$ can be written as

$$\bar{z} = \min \sum_{i=1}^{N} \sqrt{y_i}$$

subject to

$$\sum_{i=1}^{N} y_i = \sigma^2$$

$$y_i + s_i = \bar{\sigma}^2 \qquad i = 1, 2, \ldots, N$$

$$y_i - t_i = \underline{\sigma}^2 \qquad i = 1, 2, \ldots, N$$

$$y_i,\ s_i,\ t_i \geqslant 0 \qquad i = 1, 2, \ldots, N.$$

The objective function is strictly concave and the feasible set is compact. Hence, there is a unique optimal solution which is an extreme point of $F$, the feasible set of $(W)$.

Every extreme point of $F$ has $2N + 1$ basic variables. Assume that $y_{i1}$ and $y_{i2}$ are in the open interval $(\underline{\sigma}^2, \bar{\sigma}^2)$. Then, the corresponding slacks $s_{i1}$, $s_{i2}$, $t_{i1}$ and $t_{i2}$ are strictly positive. We are left with $2N + 1 - 6 = 2(N - 2) - 1$ variables in the basis. The remaining $2(N - 2)$ bounding constraints must have each at least one variable in the basis. However, this is impossible since we are left with $2(N - 2) - 1$ basic variables. Consequently, at most, one variable $y_i$ can be in the open interval $(\underline{\sigma}^2, \bar{\sigma}^2)$ for each extreme point of $F$. Thus, at every extreme point of $F$, $y_i = \underline{\sigma}^2$ or $y_i = \bar{\sigma}^2$ for all except at most one $i$. Therefore, the minimum cost solution will have as many variables equal to $\underline{\sigma}^2$ as possible while still satisfying the knapsack constraint $\sum_{i=1}^{N} y_i = \sigma^2$.

The optimal value of $(N)$ is of the form

$N_1$ components equal to $\underline{\sigma}^2$,

$N_2$ components equal to $\bar{\sigma}^2$ and

$\gamma$ components in the open interval $(\underline{\sigma}^2, \bar{\sigma}^2)$

where $N_1$, $N_2$ are nonnegative integers, $\gamma$ is either zero or one, and $N = N_1 + N_2 + \gamma$.

From the knapsack constraint it follows that

$$N_1\underline{\sigma}^2 + N_2\bar{\sigma}^2 + \gamma\tau = \sigma^2$$

where $\tau \in (\underline{\sigma}^2, \overline{\sigma}^2)$. Thus,

$$N_1\underline{\sigma}^2 + (N_2 + \gamma)\overline{\sigma}^2 \geqslant \sigma^2$$

$$N_1\underline{\sigma}^2 + (N - N_1)\overline{\sigma}^2 \geqslant \sigma^2$$

and

$$N_1 \leqslant \frac{N\overline{\sigma}^2 - \sigma^2}{\overline{\sigma}^2 - \underline{\sigma}^2} \tag{3.1}$$

$\overline{z}$ can be written as

$$\overline{z} = N_1\underline{\sigma} + N_2\overline{\sigma} + \gamma\tau.$$

Consequently, $N_1$ should be as large as possible in the optimal solution. From (3.1),

$$N_1 = \left\lceil \frac{N\overline{\sigma}^2 - \sigma^2}{\overline{\sigma}^2 - \sigma^2} \right\rceil.$$

Note that if $\dfrac{\sigma^2 - N_1\underline{\sigma}^2}{\overline{\sigma}^2}$ is integer then $\gamma = 0$. Otherwise $\gamma = 1$. Therefore,

$$\gamma = \left\lfloor \frac{\sigma^2 - N_1\underline{\sigma}^2}{\overline{\sigma}^2} - \left\lceil \frac{\sigma^2 - N_1\underline{\sigma}^2}{\overline{\sigma}^2} \right\rceil \right\rfloor \qquad \square.$$

Let $k = \dfrac{c_2}{c_1 + c_2}$ and let $\zeta$ be a normally distributed random variable with mean zero and standard deviation equal to one. Denote by $\zeta_k$ the value of $\zeta$ such that

$$\text{Prob}\,(\zeta \leqslant \zeta_k) = k.$$

It is well known [6] that the optimal lot size for item $i$ is

$$Q_i^* = \mu_i + \zeta_k\sigma_i. \tag{3.2}$$

The total under-stock and over-stock cost for item $i$ is

$$TC_i = co_i(Q^* - \mu_i) + (co_i + cu_i)\,\sigma_i G(\zeta_k) \tag{3.3}$$

where

$$G(\zeta_k) = \int_{\zeta_k}^{\infty} (\zeta_k - \zeta) f(\zeta)\, d\zeta$$

and $f(\cdot)$ is the normal probability density function [6].

Substituting (3.2) in (3.3)

$$TC_i = co_i\zeta_k\sigma_i + (co_i + cu_i)\sigma_i G(\zeta_k) = [c_1\zeta_k + (c_1 + c_2)G(\zeta_k)]\mathrm{v}_i\sigma_i.$$

The total cost of under-stock and over-stock for all items if each is ordered separately is

$$TC = \sum_{i=1}^{N} TC_i = [c_1\zeta_k + (c_1 + c_2)\,G(\zeta_k)] \sum_{i=1}^{N} \mathrm{v}_i\sigma_i.$$

Let

$$M = c_1\zeta_k + (c_1 + c_2)\,G(\zeta_k).$$

Hence,

$$TC = M\sum_{i=1}^{N} v_i\sigma_i. \tag{3.4}$$

Assume the $N$ items are aggregated in a single product with parameters $co = c_1 v$, $cu = c_2 v$, $v = \sum_{i=1}^{N} v_i$, $\mu = \sum_{i=1}^{N} \mu_i$ and $\sigma^2 = \sum_{i=1}^{N} \sigma_i^2$. The demand of the aggregate product, or, equivalently, the family, is single period and is normally distributed with mean $\mu$ and standard deviation $\sigma$.

The optimal ordering quantity and cost for the family are

$$Q_A^* = \mu + \zeta_k\sigma$$

and

$$TC_A' = M\, v\sigma.$$

Note that $TC_A'$ would be equal to $TC$ is $\sigma$ was equal to $\sum_{i=1}^{N} \sigma_i$.

We propose to use as an approximation for $TC$ the value $TC_A$ defined as

$$TC_A = TC_A' \frac{1}{\sqrt{N}} = M\, v\sigma N^{-\frac{1}{2}}.$$

The quality of the approximation can be measured by the following relative error

$$\Delta = \frac{TC_A - TC}{TC_A} = \frac{M v\sigma N^{-\frac{1}{2}} - M\sum_{i=1}^{N} v_i\sigma_i}{M v\sigma N^{-\frac{1}{2}}} = 1 - \sqrt{N}\sum_{i=1}^{N} \frac{v_i\sigma_i}{v\sigma}. \tag{3.5}$$

Given a finite sequence $\{x_i\}_i$ let $\bar{x}$ and $\underline{x}$ denote, respectively, $\max_i\{x_i\}_i$ and $\min_i\{x_i\}_i$. Define

$$a_1 = \frac{N\bar{v}}{v},\ a_2 = \frac{N\underline{v}}{v},\ \beta_1 = \frac{\sqrt{N}\,\bar{\sigma}}{v},\ \beta_2 = \frac{\sqrt{N}\,\underline{\sigma}}{\sigma}.$$

Note that $a_1$ is the quotient of the maximum and average unit value of the $N$ items. Similar interpretations hold for the other parameters. Moreover,

$$a_1 \geqslant 1,\ a_2 \leqslant 1,\ \beta_1 \geqslant 1,\ \beta_2 \leqslant 1. \tag{3.6}$$

From (3.5),

$$\Delta = 1 - \sqrt{N}\sum_{i=1}^{N} \frac{v_i\sigma_i}{v\sigma} \geqslant 1 - \frac{\sqrt{N}\,\bar{v}}{v\sigma}\sum_{i=1}^{N} \sigma_i = 1 - \frac{a_1}{\sigma\sqrt{N}}\sum_{i=1}^{N} \sigma_i$$

and by Proposition 3.1,

$$\Delta \geqslant 1 - \frac{a_1}{\sigma\sqrt{N}}\sqrt{N}\sigma = 1 - a_1. \tag{3.7}$$

From (3.5)

$$\Delta = 1 - \sqrt{N} \sum_{i=1}^{N} \frac{v_i \sigma_i}{v\sigma} \geq 1 - \frac{\sqrt{N}\,\underline{\sigma}}{\sigma} = 1 - \beta_1. \tag{3.8}$$

Upper bounds on the value of $\Delta$ can also be derived from (3.5):

$$\Delta = 1 - \sqrt{N} \sum_{i=1}^{N} \frac{v_i \sigma_i}{v\sigma} \leq 1 - \frac{\sqrt{N}\,\underline{\sigma}}{\sigma} = 1 - \beta_2. \tag{3.9}$$

Similarly,

$$\Delta = 1 - \sqrt{N} \sum_{i=1}^{N} \frac{v_i \sigma_i}{v\sigma} \leq 1 - \frac{\sqrt{N}\,\underline{v}}{v} \frac{\sum_{i=1}^{N} \sigma_i}{\sigma} = 1 - a_2 \frac{\sum_{i=1}^{N} \sigma_i}{\sqrt{N}\,\sigma}.$$

By Proposition 3.2,

$$\Delta \leq 1 - a_2 \frac{\bar{z}}{\sqrt{N}\,\sigma}. \tag{3.10}$$

The next proposition summarizes the results obtained.

PROPOSITION 3.3: The following bounds hold for

$$\Delta = \frac{TC_A - TC}{TC_A}$$

(1) $1 - a_1 \leq \Delta \leq 1 - a_2 \dfrac{\bar{z}}{\sqrt{N}\,\sigma}$

(2) $1 - \beta_1 \leq \Delta \leq 1 - \beta_2$.

The bounds on $\Delta$ depend on the extremal parameters of the items and the family parameters. These can be estimated with greater accuracy and less cost than the items parameters. Moreover, (1) and (2) in Proposition 3.3 hold with $a_1$, $a_2$, $\beta_1$, $\beta_2$ replaced by respective bounds. Therefore, the computation of $TC_A$ may play a very useful role in the diagnostic analysis of a family of items with single period stochastic demands.

## 4. CONCLUSIONS AND TOPICS FOR FUTURE RESEARCH

In this paper we have developed diagnostic models for a limited class of inventory-control problems. We have shown the effect of different aggregation schemes and their power to estimate the performance to be expected from the implementation of a recommended system. Upper and lower bounds were provided for that purpose.

A very important part of a diagnostic analysis is the evaluation of the performance of different inventory-control algorithms. We limited our analysis to consider two basic situations: a case where the items of the system share a common setup, and a case where the products have single-season-independent demands. In those situations, the following diagnostic-analysis scheme is suggested:

—Group the items into families according to some specific rules depending on the particularities of the products. Proposition 2.6 proved that similar seasonal behavior gives optimal aggregate

schedules for items sharing a set up. Proposition 3.3 proved that, for the single period case, either same item values or same standard deviation of the item demands also yield optimal aggregate models.

—Compute aggregate information on the critical parameters used to describe inventory performance.

—Solve the aggregate models, and compute lower and upper bounds for the optimal value of the total inventory cost.

—The bounds provide a range within which the inventory cost will reside if the decision rules being considered would be implemented to the actual inventory system. Thus, we have obtained a first approximation to assess the benefits to be derived from the application of those decision rules.

We believe that diagnostic analyses are important practical tools. The have been neglected in the literature and a great deal of work remains to be done to determine effective diagnostic methodologies not only for inventory systems but also for logistics problems in general. Within inventory-control systems the next natural step is to develop diagnostic models to study items that do not share the same setup, items subject to quantity discounts, items subject to inventory constraints and other cases cited in the taxonomies provided by Silver [7], Nahmias [5], and Aggarwal [1]. Production planning and distribution systems are other areas for potential development of diagnostic models. The need for such methods should also motivate researchers to explore the concepts of aggregation and heuristics not only under the view point of computational complexity but also to prevent the massive data collection and processing that would make diagnostic studies prohibitive.

## ACKNOWLEDGMENTS

The authors are grateful to Professor Thomas L. Magnanti and Dr. Debashish Sarkar for their careful reading of an earlier version of the paper.

## REFERENCES

[1] Aggarwal, S.C., "A Review of Current Inventory Theory and Its Applications," International Journal of Production Research, *12*, 443-482 (1974).

[2] Bitran, G.R., A.C. Hax and J. Valor-Sabatier, "Diagnostic Analysis of Inventory Systems: A Statistical Approach," Sloan School of Management Technical Report No. 10 (February 1981).

[3] Evans, J.R., "Aggregation in the Generalized Transportation Problem," Computers and Operations Research, *6*, 199-204 (1979).

[4] Hax, A.C., N.S. Majluf, and M. Pendrock, "Diagnostic Analysis of a Production and Distribution System," Management Science *26*, 171-189 (September 1980).

[5] Nahmias, S., "Inventory Models," The Encyclopedia of Computer Science and Technology, *9*, 447-483 (1978).

[6] Peterson, R. and E.A. Silver, Decision Systems for Inventory Management and Production Planning (John Wiley and Sons, New York, 1979).

[7] Silver, E.A., "Operations Research in Inventory Management. A Review and Critique," Operations Research, *29*, 628-645 (July-August 1981).

[8] Veinott, A.F. Jr., "Extreme Points of Leontief Substitution Systems," Linear Algebra and its Applications, *1*, 181-194 (1968).

[9] Wagner, H.M., "Research Portfolio for Inventory Management and Production Planning Systems," Operations Research, *28*, 445-475 (May-June 1980).

[10] Wagner, H.M. and T.M. Whitin, "Dynamic Version of the Economic Lot Size Model," Management Science, *5*, 89-96 (October 1958).

[11] Winters, P.R., "Forecasting Sales by Exponentially Weighted Moving Averages," Management Science, *6*, 324-342 (1960).
[12] Zipkin, P.H., "Bounds on the Aggregation of Linear Programs," Operations Research, *28*, 403-418 (March-April 1980).

# PRODUCTION PLANNING FOR MULTI-RESOURCE NETWORK SYSTEMS*

Robert C. Leachman

*University of California, Berkeley*
*Berkeley, California*

ABSTRACT

Production planning for large-scale production systems requiring the allocation of numerous resources is considered. It is demonstrated how the dynamic activity analysis developed by Shephard leads to linear programming solutions of production planning problems. Three types of planning problems are formulated: maximization of output levels for a given time horizon; minimization of production duration for given output histories; and minimization of production costs for given output histories.

## 1. INTRODUCTION

Previous efforts (von Lanzenauer [4] and Candea [1]) to mathematically model capacitated, multistage production systems have been motivated by manufacturing shop environments, in which many products are to be produced using a given network of facilities.† The problem considered is to determine workforce levels and product lot sizes in each time period so as to minimize costs to meet external demand schedules [1] [4].

The focus of this paper is planning for production systems in which the production network elements are dedicated to producing a single product, but allocation of numerous resources among the elements is required, and other kinds of production planning problems are posed. Shephard, Al-Ayat and Leachman [7], and Shephard and Al-Ayat [8] have developed a continuous flow dynamic activity analysis model of production, in which a network of activities characterizes the component tasks of production. Required facilities and other resources are considered as inputs to be allocated among the activities. The presentation in [8] is taken as an appropriate point of departure here. In the following, this model is extended to include inventory capacities as in [4], initial inventories of intermediate products, and classes of exogenous inputs.

Three types of production planning problems are formulated and solved using linear programming methods. The problem types considered are maximization of output accumulations by a given horizon; minimization of production duration for required output histories; and minimization of costs for given output requirements.

## 2. THE MODEL

Following Shephard and Al-Ayat [8], the production system is viewed as a network of production activities which are denoted by $A_1, A_2, \ldots, A_N$. In the network, nodes represent activities and arcs

*This research was supported by the Office of Naval Research under Contract N00014-76-C-0134 with the University of California.

†No attempt is made here to reference multistage modeling efforts of uncapacitated systems or pure serial or parallel structures; for a survey of such efforts, see Candea [1].

indicate intermediate product transfers, i.e., the use of each activity's output as input by other activities. The operation of each activity $A_i$ is measured in terms of an intensity function $z_i(t)$, $t = 0, 1, 2, \ldots$, whose value at time $t$ indicates activity input during $[t, t+1)$ and output at time $(t+1)$ when taken with technical coefficients defined as follows:

(a) $c_i(t)$, $t=0, 1, 2, \ldots, i=1, \ldots, N$, where $c_i(t)$ is the amount of output of activity $A_i$ at time $t$ per unit intensity of activity $A_i$.

(b) $a_{ik}(t)$, $t=0, 1, 2, \ldots, k=1, 2, \ldots, NK$, $i=1, \ldots, N$, where $a_{ik}(t)$ is the amount of exogenous input type $k$ required at time $t$ per unit intensity of activity $A_i$. The first $NS \leqslant NK$ inputs are designated nonstorable resources which cannot be accumulated; the remaining exogenous inputs can be accumulated, and are termed storable resources.

(c) $\bar{a}_{ij}(t)$, $t=0, 1, 2, \ldots, i=1, \ldots, N$, $j=1, \ldots, N$, where $\bar{a}_{ij}(t)$ is the amount of intermediate product from activity $A_i$ required at time $t$ per unit intensity of operating $A_j$.

For a time horizon $T$ for production activity, we introduce the following technical limitations on the system:

(a) $\{\bar{z}_i(t)\}_{i=1,\ldots,N}^{t=0,1,2,\ldots,T-1}$, the activity intensity bounds, natural bounds resulting from available workspace and other limitations not considered as exogenous input;

(b) $\{X_k(t)\}_{k=1,\ldots,NS}^{t=0,1,2,\ldots,T-1}$, the time histories of nonstorable resource levels available for input to the system;

(c) $\{Y_k(t)\}_{k=NS+1,\ldots,NK}^{t=0,1,\ldots,T-1}$, the time histories of storable resources made available to the system, where

$$\sum_{\tau=0}^{t} Y_k(\tau)$$

is the cumulative amount of resource $k$ supplied during $[0, t)$;

(d) $\{\text{inv}_i^0\}_{i=1,\ldots,N}$, the initial inventories of activity product for intermediate uses; and

(e) $\{\text{cap}_i(t)\}_{i=1,\ldots,N}^{t=0,1,\ldots,T-1}$, the bounds on accumulations of activity product awaiting intermediate uses, arising from limited storage capacity for in-process inventories.

For our purposes here, the intensity $z_i(t)$ of activity $A_i$, $i=1, \ldots, N$, on each time interval $[t, t+1)$, $t=0, 1, 2, \ldots, T-1$, shall be partitioned into effort producing intermediate product, $z_i^I(t)$, and effort producing final product, $z_i^F(t)$, where

$$z_i^I(t) + z_i^F(t) = z_i(t). \tag{1}$$

These variables indicate the allocation of activity output produced during $[t, t-1)$ to final and intermediate uses.

A production plan is a specification over some finite period $[0, T)$ of the activity intensities

$$\{z_i^I(t),\ z_i^F(t)\}_{i=1,\ldots,N}^{t=0,\ldots,T-1}.$$

Such a plan is said to be *feasible for* $L(T)$ if the plan belongs to the set $\mathbf{L(T)}$ defined by the following inequalities:

$$\mathbf{L(T)1.} \qquad \sum_{i=1}^{N} a_{ik}(t)\,[z_i^I(t) + z_i^F(t)] \leqslant X_k(t),\ k = 1, \ldots, NS,\ t = 0, 1, \ldots, T-1.$$

$$\mathbf{L(T)2.} \qquad \sum_{\tau=0}^{t} \sum_{i=1}^{N} a_{ik}(\tau)\,[z_i^I(\tau) + z_i^F(\tau)] \leqslant \sum_{\tau=0}^{t} Y_k(\tau),\ k = NS+1, \ldots, NK,$$

$$t = 0, \ldots, T-1.$$

$$\mathbf{L(T)3.} \qquad \sum_{\tau=1}^{t} \sum_{i=1}^{N} \bar{a}_{ji}(\tau)\,[z_i^I(\tau) + z_i^F(\tau)] - \sum_{\tau=0}^{t-1} c_j(\tau+1) z_j^I(\tau) \leqslant \mathrm{inv}_j^0,$$

$$j = 1, \ldots, N,\ t = 1, \ldots, T-1, \text{ and}$$

$$\sum_{i=1}^{N} \bar{a}_{ji}(0)\,[z_i^I(0) + z_i^F(0)] \leqslant \mathrm{inv}_j^0,\ j = 1, \ldots, N.$$

$$\mathbf{L(T)4.} \qquad \sum_{\tau=0}^{t-1} c_j(\tau+1) z_j^I(\tau) - \sum_{\tau=1}^{t} \sum_{i=1}^{N} \bar{a}_{ji}(\tau)\,[z_i^I(\tau) + z_i^F(\tau)] \leqslant \mathrm{cap}_j(t) - \mathrm{inv}_j^0,$$

$$j = 1, \ldots, N,\ t = 1, \ldots, T-1, \text{ and}$$

$$-\sum_{i=1}^{N} \bar{a}_{ji}(0)\,[z_i^I(0) + z_i^F(0)] \leqslant \mathrm{cap}_j(0) - \mathrm{inv}_j^0,\ j = 1, \ldots, N.$$

$$\mathbf{L(T)5.} \qquad z_j^I(t) + z_j^F(t) \leqslant \bar{z}_j(t),\ j = 1, \ldots, N,\ t = 0, \ldots, T-1,$$

$$z_j^I(t),\ z_j^F(t) \geqslant 0,\ j = 1, \ldots, N,\ t = 0, \ldots, T-1.$$

Constraints **L(T)1** and **L(T)2** express resource limitations. Constraints **L(T)3** insure adequate intermediate product transfers occur to support production activity, while constraints **L(T)4** insure that inventories of intermediate products do not exceed capacities. Finally, constraints **L(T)5** limit intensities to nonnegative values less than intensity bounds.

The set of linear inequalities **L(T)** constitutes a continuous flow model of production, in which any positive intensity of activity operation supplies completed product to final or intermediate uses, or to inventory. In the case that intermediate products of a system are large, discrete units, precedence relations occur between activities output unit by output unit, and constraints **L(T)3** and **L(T)4** must be modified. This case will not be treated here, and the reader is referred to [6], in which a dynamic activity analysis was developed on a critical path analysis network.

In the case more than one activity produces a certain product, constraints **L(T)3** and **L(T)4** must be modified for the activities in question. See [6] and [8]. However, with such revisions **L(T)** still constitutes a set of linear inequalities. For simplicity of exposition, we assume in what follows that no two activities produce the same product.

## 3. PRODUCTION PROGRAMMING

### 3.1. Output Maximization

In this section, programs are formulated for the maximization of value or mix functions of final output accumulations. We consider first the case where a specific product mix of final output is desired, and the problem is to maximize the scale of this mix accumulated by a time horizon $T$.

Let $z_{N+1}$ be a variable indexing the scale of the accumulation by time $T$. The amounts of the various products will be related by coefficients

$$\bar{a}_{i,N+1},\ i = 1, \ldots, N,$$

where $\bar{a}_{i,N+1}\, z_{N+1}$ is the amount of final product from activity $A_i$ accumulated. The problem in question is formulated as a linear program as follows:

Maximize $z_{N+1}$

subject to

$$\mathbf{O}1.\quad \bar{a}_{i,N+1} z_{N+1} \leqslant \sum_{t=0}^{T-1} c_i(t+1) z_i^F(t),\ i = 1, \ldots, N.$$

$$\mathbf{O}2.\quad \{z_i^I(t), z_i^F(t)\}_{i=1,\ldots,N}^{t=0,\ldots,T-1} \in L(T),$$

$$\mathbf{O}3.\quad z_{N+1} \geqslant 0.$$

In general, the program involves $[5(N) + NK](T) + N + 1$ variables, and $[3(N) + NK](T) + N$ constraints. Clearly, the time horizon (i.e., the number of time periods) is the most sensitive factor in terms of problem size which can be handled. The structure of the constraint set can be modified by rewriting constraints L(T)3 and L(T)4 in terms of intermediate product inventory variables

$$\{\mathrm{inv}_j^I(t)\},\ t = 0, \ldots, T-1,$$

which are the slack variables defined by constraints L(T)3. Using these variables, we reformulate constraints L(T)3 and L(T)4 as follows:

$$\mathbf{L}(\mathbf{T})3.\quad \sum_{i=1}^{N} \bar{a}_{ji}(t)[z_i^I(t) + z_i^F(t)] - c_j(t) z_j^I(t-1) - \mathrm{inv}_j^I(t-1) + \mathrm{inv}_j^I(t) = 0,$$

$$j = 1, \ldots, N,\ t = 1, \ldots, T-1, \text{ and}$$

$$\sum_{i=1}^{N} \bar{a}_{ji}(0)[z_i^I(0) + z_i^F(0)] + \mathrm{inv}_j^I(0) = \mathrm{inv}_j^0,\ j = 1, \ldots, N.$$

$$\mathbf{L}(\mathbf{T})4.\quad \mathrm{inv}_j^I(t) \leqslant \mathrm{cap}_j(t),\ j = 1, \ldots, N,\ t = 0, \ldots, T-1.$$

With this revision, it is evident that the constraints L(T)1, L(T)4 and L(T)5 apply only time period by time period, and the constraint matrix exhibits partial block diagonal structure. Potential is thus offered for application of large-scale programming procedures such as decomposition. (See [5].)

We next consider the case where the value of output produced is to be maximized. We suppose each product $i$ has a constant unit price $p_i$. The maximum value of output accumulated from production activity during $[0, T)$ is then given by the optimum of the following linear program.

Maximize $$\sum_{t=0}^{T-1} \sum_{i=1}^{N} p_i c_i(t+1) z_i^F(t)$$

Subject to $$\{z_i^I(t), z_i^F(t)\}_{i=1,\ldots,N}^{t=0,\ldots,T-1} \in L(T).$$

The remarks about problem size and structure concerning the previous program apply here as well, as only the product mix variable and the $N$ constraints O1 have been deleted.

### 3.2. Time Minimization

In this kind of planning problem, there are final output demands which must be met, but the overall production duration is to be minimized. Final output demands are expressed in cumulative terms as follows. Let

$$\hat{U}_i(t),\ t = 1, 2, \ldots, T,\ i = 1, \ldots, N,$$

denote the required cumulative delivery of final product $i$ by time $t$. Here, we are considering the situation where early delivery of final products is acceptable or even desirable. These demands imply constraints

$$\sum_{\tau=0}^{t-1} c_i(\tau+1) z_i^F(\tau) \geqslant \hat{U}_i(t),\ i = 1, \ldots, N,\ t = 1, \ldots, T. \tag{2}$$

We first consider the problem of finding the latest starting time for production activity sufficient to satisfy (2). A feasible production plan for this problem would satisfy the linear inequalities (2) and $\mathbf{L(T)}$. An optimal plan would have the characteristic that

$$z_i(0) = z_i(1) = z_i(2) = \cdots = z_i(t_0) = 0,\ i = 1, \ldots, N, \tag{3}$$

where $t_0$ is as large as possible. Such a plan may be found (if one exists) by solving a sequence of Phase I linear programs (see [2]) as follows.

The set of inequalities under consideration is of the form

$$\mathbf{Az} + \mathbf{Bx} = \mathbf{b} \tag{4}$$
$$\mathbf{z} \geqslant 0,\ \mathbf{x} \geqslant 0,$$

where

$$\mathbf{z} = (z_1^I(0), z_1^F(0), \ldots, z_N^I(0), z_N^F(0), z_1^I(1), z_1^F(1), \ldots, z_N^I(1), z_N^F(1), \ldots, z_1^I(T-1), z_1^F(T-1), \ldots, z_N^I(T-1), z_N^F(T-1));$$
$$\mathbf{x} = (x_1, \ldots, x_{m+n});$$
$$m = (N)(T),\ n = (NK)(T) + 3(N)(T),\ l = 2(N)(T);$$

**A** is the $(m+n) \times l$ matrix of coefficients of activity intensities in (2) and $\mathbf{L(T)}$, where the first $m$ rows arise from (2);

**B** is the $(m+n) \times (m+n)$ matrix of coefficients of slack variables for said constraints; and

**b** is the $(m+n)$ vector of right hand side constant terms of the constraints.

For the inequalities organized in this fashion, the solution algorithm is presented below:

STEP 0: Initialize $\tau = T - 1$.

STEP 1: Solve the Phase I problem with the first $(N)(\tau)$ columns of the tableau corresponding to (4) deleted. If a feasible solution is found, stop; then

$$\{z_i^I(t),\ z_i^F(t)\}_{i=1,\ldots,N}^{t=0,\ldots,T-1}$$

is an optimal production plan. If the problem is infeasible, go to Step 2.

STEP 2. If $\tau = 0$, stop; then the set of inequalities is infeasible. Otherwise, decrease $\tau$ to $\tau - 1$ and go to Step 1.

The algorithm is seen to initially ignore all columns in the tableau corresponding to activity intensities in periods before time $(T-1)$, and to then attempt to find a basic solution. If none can be found, columns corresponding to

$$z_i^I(T-2),\ z_i^F(T-2),\ i = 1, \ldots, N,$$

are also considered. The algorithm continues to allow the use of columns corresponding to activity operation one time period before the earliest period of activity operations allowed by the previous iteration. The algorithm terminates either the first time a feasible basic solution is found, or else all columns are adjoined without finding one. In the former case, an optimal production plan is found, and in the latter case, the output schedule (2) is infeasible for the limitations $\mathbf{L}(\mathbf{T})$.

We next consider the problem of finding the earliest time all product accumulations can be completed. It is immediately apparent that an approach similar to that considered above can be used to solve this problem. A sequence of Phase I procedures is again suggested, but in this case starting with the possibility of positive activity intensities only during $[0, 1)$, and proceeding forwards in time. Later production activity is allowed period by period until either a feasible basic solution is found for the first time, or else the horizon is reached without finding one. In the former case, an optimal production plan is obtained, and in the latter case, the output schedule is infeasible for the limitations imposed.

### 3.3. Cost Minimization

In this section we formulate the problem of determining a minimum cost production plan which meets a given final output schedule expressed in the form of (2).

Nonstorable resources are assumed to have capacity costs corresponding to the peak demands for each such resource. These resources cannot be accumulated, so that the production system must have the capability to accommodate peak loads. Storable resources, however, have *prices*; these resources account for the variable cost of production activity. It is assumed that storable resources can be procured as required, so that inventories of same are ignored. We assume intermediate product inventories also have capacity costs, corresponding to peak storage requirements. These inventories will also bear holding costs representing opportunity charges for unproductive capital.

Assuming linear capacity costs, the problem is formulated as a linear program as follows. Let

$$\mathbf{C}^x(t) = (C_1^x(t), \ldots, C_{NS}^x(t)),$$

be the vector of costs per unit capacity for nonstorable resources maintained during $[t, t+1)$; let

$$\mathbf{C}^y(t) = (C_{NS+1}^y(t), \ldots, C_N^y(t)),$$

be the price vector for storable resources procured for use during $[t, t+1)$; let

$$\mathbf{C}^z(t) = (C_1^z(t), \ldots, C_N^z(t)),$$

be the vector of costs per unit storage capacity maintained during $[t, t+1)$ for intermediate products; and let

$$\mathbf{H}(t) = (H_1(t), \ldots, H_N(t)),$$

be the holding costs for intermediate products held during $[t, t+1)$.

To serve as variables in the minimization, let

$$\mathbf{X} = (X_1, \ldots, X_{NS}),$$

denote the peak requirements in any unit time interval of nonstorable resources; let

$$\mathbf{cap} = (\mathrm{cap}_1, \ldots, \mathrm{cap}_N),$$

denote the required intermediate product inventory storage capacities; and let

$$\{\mathrm{inv}_i^I(t), z_i^I(t), z_i^F(t)\}_{i=1,\ldots,N}^{t=0,\ldots,T-1}$$

be the inventory and intensity variables as before.

For given intensity bounds

$$\{\bar{z}_i(t)\}_{i=1,\ldots,N}^{t=0,\ldots,T-1}$$

and initial intermediate product inventories

$$\{\text{inv}_j^0\}_{j=1,\ldots,N},$$

the minimum cost production plan meeting the output schedule (2) is given by the optimum of the linear program

$$\text{Minimize} \quad \sum_{t=0}^{T-1} \mathbf{C}^x(t) \cdot \mathbf{X} + \sum_{t=0}^{T-1} \sum_{k=NS+1}^{NK} \sum_{i=1}^{N} C_k^y(t)\, a_{ik}(t)\,[z_i^I(t) + z_i^F(t)]$$
$$+ \sum_{t=0}^{T-1} \mathbf{C}^z(t) \cdot \mathbf{cap} + \sum_{t=0}^{T-1} \sum_{i=1}^{N} H_i(t)\, \text{inv}_i^I(t)$$

subject to

C1. $$\sum_{\tau=0}^{t-1} c_i(\tau+1) z_i^F(\tau) \geqslant \hat{U}_i(t), \quad i=1, \ldots, N, \quad t=1, \ldots, T.$$

C2. $$\sum_{i=1}^{N} a_{ik}(t)\,[z_i^I(t) + z_i^F(t)] - X_k \leqslant 0, \quad k=1, \ldots, NS, \quad t=0, 1, \ldots, T-1.$$

C3. $$\sum_{i=1}^{N} \bar{a}_{ji}(t)\,[z_i^I(t) + z_i^F(t)] - c_j(t) z_j^I(t-1) - \text{inv}_j^I(t-1) + \text{inv}_j^I(t) = 0,$$

$$j = 1, \ldots, N, \quad t=1, \ldots, T-1, \text{ and}$$

$$\sum_{i=1}^{N} \bar{a}_{ji}(0)\,[z_i^I(0) + z_i^F(0)] + \text{inv}_j^I(0) = \text{inv}_j^0, \quad j=1, \ldots, N.$$

C4. $$\text{inv}_j^I(t) - \text{cap}_j(t) \leqslant 0, \quad j=1, \ldots, N, \quad t=0, \ldots, t-1.$$

C5. $$z_j^I(t) + z_j^F(t) \leqslant \bar{z}_j(t), \quad j=1, \ldots, N, \quad t=0, \ldots, T-1.$$

C6. $$\mathbf{X} = (X_1, \ldots, X_{NS}) \geqslant 0,$$

$$\mathbf{cap} = (\text{cap}_1, \ldots, \text{cap}_N) \geqslant 0,$$

$$\mathbf{inv}^I(t) = (\text{inv}_1^I(t), \ldots, \text{inv}_N^I(t)) \geqslant 0, \quad t=0, \ldots, T-1,$$

$$z_j^I(t),\ z_j^F(t) \geqslant 0, \quad j=1, \ldots, \mathrm{N}, \quad t=0, \ldots, T-1.$$

Here constraints C2 define the required nonstorable resource capacities, and constraints C4 define the required intermediate product storage capacities. Constraints C3 and C5 deal with inventory balance and intensity bounds in the same manner as the treatment of output maximization problems, while constraints C1 repeat (2).

In general, the program includes $(4N + NS)(T+1)$ variables and $(4N + NS)(T)$ constraints. As before, the fineness of the time grid is the most sensitive factor in terms of the problem size which can in practice be solved. A bordered angular configuration for the constraint matrix is now displayed, in which constraints C2, C4, and C5 exhibit a block diagonal structure with coupling variables. Although this is a more difficult structure than that for the output maximization problems, nonetheless it can be exploited. See [5].

As an alternative to the constant capacities for each nonstorable resource defined by constraints C2, one may allow capacities to be adjusted from time period to time period according to linear costs. Many authors have formulated labor workforce levels in this fashion, allowing hiring and firing in each period. See for example [3] or [4]. Such formulations may be integrated here as appropriate.

## REFERENCES

[1] Candea, D., "Issues of Hierarchical Planning in Multi-Stage Production Systems," Technical Report No. 134, Operations Research Center, Massachusetts Institute of Technology (1977).

[2] Gass, S.I., *Linear Programming — Methods and Applications* (McGraw-Hill, New York, 1975).

[3] Johnson, L.A. and D.C. Montgomery, *Operations Research in Production Planning, Scheduling, and Inventory Control* (John Wiley & Sons, New York, 1970).

[4] von Lanzenauer, C.H., "Production and Employment Scheduling in Multistage Production Systems," Naval Research Logistics Quarterly, *17*, No. 2 (1970).

[5] Lasdon, L.S., *Optimization Theory For Large Systems* (MacMillan Co., New York, 1970).

[6] Leachman, R.C., "Aspects of Dynamic Production Planning," Report ORC 79-8, Operations Research Center, University of California, Berkeley (1979).

[7] Shephard, R.W., R. Al-Ayat, and R.C. Leachman, "Shipbuilding Production Function: An Example of a Dynamic Production Function," *Quantitative Wirtschaftsforschung Festchrift Volume (Wilhelm Krelle),* J.B.C. Mohr, Editor (Tubingen, West Germany 1977).

[8] Shephard, R.W. and R. Al-Ayat, "Dynamic Theory of Production Correspondences, Part IV," Report ORC 78-15, Operations Research Center, University of California, Berkeley (1978), to appear in revised form as "Dynamic Theory of Production Networks," *Mathematical Systems in Economics* (Anton Hain, Meisenheim am Glan, West Germany).

# THE CLASS OF MIFRA LIFETIMES AND ITS RELATION TO OTHER CLASSES*

Henry W. Block and Thomas H. Savits

*Department of Mathematics and Statistics*
*University of Pittsburgh*
*Pittsburgh, Pennsylvania*

ABSTRACT

In a previous paper, the authors have introduced a class of multivariate lifetimes (MIFRA) which generalize the univariate lifetimes with increasing failure rate average (IFRA). They have also shown that this class satisfies many fundamental properties. In this paper it is shown thai other concepts of multivariate IFRA do not satisfy all of these properties. Relationships between MIFRA and these other concepts are given. Finally, positive dependence implications with respect to these classes are also discussed.

## 1. INTRODUCTION

The class of univariate lifetimes with increasing failure rate average (IFRA) has been of great importance in reliability theory. The importance of the class, and properties thereof, are discussed in the text of Barlow and Proschan [1] whose notation and terminology are followed here. A recent development with respect to this class has been the resolution, by Block and Savits [2], of a long standing problem concerning the closure of this class under convolution.

Several recent papers by Block and Savits [3, 4], and by Esary and Marshall [6], have proposed various multivariate extensions of this univariate class. It is our purpose in the present paper to give the relations among these various concepts and to show that one of these concepts, which was designated MIFRA in Block and Savits [3], is preferable to others. This will be done by showing that among these various extensions only the MIFRA class of distribution satisfies all of the properties which one would reasonably expect for a class of multivariate IFRA distributions. Furthermore, dependence properties and the lack thereof for those classes are also discussed.

One deviation which we shall make from the notation of Barlow and Proschan [1] is to call a structure function $\phi(\underline{x})$ *monotone* if it is increasing in its components and in addition $\phi(\underline{0}) = 0$ and $\phi(\underline{1}) = 1$. Esary and Marshall [5] have called such a function coherent. We conform to the terminology of Barlow and Proschan [1], and call a structure function *coherent* (called fully coherent by Esary and Marshall) if it is increasing in its arguments and if all components are essential. The life function $\tau$ corresponding to a system $\phi$ is called *monotone* (*coherent*) if $\phi$ is monotone (coherent). See Esary and Marshall [5] for a discussion of life functions.

## 2. MULTIVARIATE IFRA

Block and Savits [3] have introduced a concept of multivariate IFRA which is given in the following definition.

*This research was supported by ONR Contract N00014-76-C-0839 and by NSF Grant MC577-01458.

DEFINITION 1. Let $\underline{T} = (T_1, \ldots, T_m)$ be a nonnegative random vector. The vector $\underline{T}$ is said to be *MIFRA* iff

$$E^{\alpha}[h(\underline{T})] \leqslant E[h^{\alpha}(\underline{T}/\alpha)]$$

for all continuous nonnegative increasing functions $h$ and all $0 < \alpha \leqslant 1$.

The above definition can be better appreciated by first examining a special case. If we let $h(\underline{T}) = I_A(\underline{T})$ where $A = (t_1, \infty) \times (t_2, \infty) \times \ldots \times (t_m, \infty)$ and $I$ is the indicator function of the set $A$ (i.e. $I_A(\underline{T}) = 1 \Leftrightarrow \underline{T} \in A$), the defining condition reduces to $\bar{F}^{\alpha}(\underline{t}) \leqslant \bar{F}(\alpha \underline{t})$ for all $0 < \alpha \leqslant 1$ and all $\underline{0} \leqslant \underline{t}$ where $\bar{F}(\underline{t}) = P(\underline{T} > \underline{t})$. This is a simple multivariate analog of the usual univariate definition of IFRA (see *Remark* on p. 84 of [1]) and is given in $A$ below. Furthermore, monotone systems have lifetimes of the form $\tau_1(\underline{T}) = \max_{1 \leqslant j \leqslant p} \min_{i \in P_j} T_i$ and $\tau_2(\underline{T}) = \min_{1 \leqslant j \leqslant k} \max_{i \in K_j} T_i$ where $P_j$ are min path sets and the $K_j$ min cut sets. These functions have the property that $\tau(\alpha \underline{T}) = \alpha \tau(\underline{T})$ for every $0 \leqslant \alpha \leqslant 1$ and are called homogeneous functions. Marshall and Shaked [8] have pointed out that Definition 1 is equivalent to $g(\underline{T})$ being IFRA for every increasing homogeneous function $g$. This implies that if $\underline{T}$ is MIFRA then every monotone system formed using $\underline{T}$ is IFRA.

Several other possible conditions for multivariate IFRA have been proposed.

DEFINITION 2. Let $\underline{T} = (T_1, \ldots, T_m)$ be a nonnegative random vector with survival function $\bar{F}(\underline{t}) = P(\underline{T} > \underline{t})$. The vector $\underline{T}$ is said to satisfy condition $i$ if the condition following $i$ is satisfied where $i = A, B, C, \Sigma, D, E$, or $F$.

$A$: $\bar{F}^{\alpha}(\underline{t}) \leqslant \bar{F}(\alpha \underline{t})$ for all $0 < \alpha \leqslant 1$ and all $\underline{0} \leqslant \underline{t}$.

$B$: $\underline{T}$ is such that each monotone system formed from $\underline{T}$ is univariate IFRA.

$C$: $\underline{T}$ is such that there exist independent IFRA random variables $X_1, \ldots, X_k$ and monotone life functions $\tau_i$, $i = 1, \ldots, m$ such that $T_i = \tau_i(X_1, \ldots, X_k)$ for $i = 1, \ldots, m$.

$\Sigma$: $\underline{T}$ is such that there exist independent IFRA random variables $X_1, \ldots, X_k$ and nonempty sets $S_i$ of $\{1, \ldots, k\}$ such that $T_i = \sum_{j \in S_i} X_j$ for $i = 1, \ldots, m$.

$D$: $\underline{T}$ is such that there exist independent IFRA random variables $X_1, \ldots, X_k$ and nonempty subsets $S_i$ of $\{1, \ldots, k\}$ such that $T_i = \min_{j \in S_i} X_j$ for $i = 1, \ldots, m$.

$E$: $\underline{T}$ is such that the minimum of any subfamily of $T_1, \ldots, T_m$ is IFRA.

$F$: $\underline{T}$ is such that $\min_i a_i T_i$ is IFRA for all $a_i \geqslant 0$, $i = 1, \ldots, m$.

Conditions $A$, $B$, $C$, $D$, $E$, $F$ have been given by Esary and Marshall [6] and condition $\Sigma$ was given by Block and Savits [4].

## 3. RELATIONSHIPS AMONG THE CONDITIONS

The following relationships hold between MIFRA and the seven conditions given in Section 2.

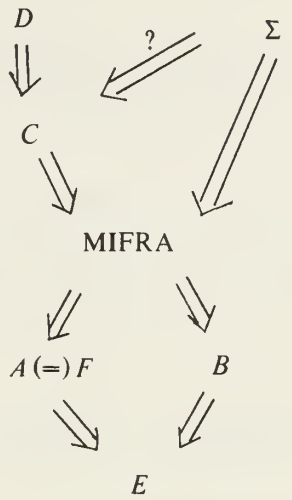



FIGURE 1. Relationships among the conditions

With the exception of the implication $\Sigma \Rightarrow C$, the above figure is complete, i.e. no more implications are possible. It is not known where $\Sigma \Rightarrow C$ holds, but we conjecture that it does not. Proofs of the remaining implications and counterexamples will now be given.

Because of results in Esary and Marshall [6] we need only show how MIFRA and $\Sigma$ compare with concepts $A$, $B$, $C$, $D$, $E$ and with each other.

### 3.1 Comparison of MIFRA with Other Conditions

a. $C \Rightarrow$ MIFRA. This follows from (iii) of Theorem 4.1 of Block and Savits [3].

b. $\Sigma \Rightarrow$ MIFRA. See (iv) of Theorem 4.1, [3].

c. MIFRA $\Rightarrow F$. Apply (P1) and (P5) of Theorem 2.3, [3].

d. MIFRA $\Rightarrow B$. This is (P1) of Theorem 2.3, [3].

e. MIFRA $\not\Rightarrow \Sigma$. Given in Example 3.3 of Block and Savits [4].

f. MIFRA $\not\Rightarrow C$. (and MIFRA $\not\Rightarrow D$). Example 3.2, [4].

g. $A \not\Rightarrow$ MIFRA. Since $A \not\Rightarrow B$.

h. $B \not\Rightarrow$ MIFRA. Since $B \not\Rightarrow A$.

### 3.2 Comparison of $\Sigma$ with Other Conditions

a. $\underline{D \not\Rightarrow \Sigma}$. Example 3.3 of Block and Savits [4].

b. $\underline{C \not\Rightarrow \Sigma}$. Since $D \not\Rightarrow \Sigma$.

c. $\underline{\Sigma \not\Rightarrow D}$. Let $X_1, X_2, X_3$ be absolutely continuous IFRA random variables. Form $Y_1 = X_1 + X_3$ and $Y_2 = X_2 + X_3$. By definition $(Y_1, Y_2)$ satisfies $\Sigma$, but by Section 10 of Esary and Marshall [6] $(Y_1, Y_2)$ does not satisfy $D$.

All other counterexamples and implications, with the exception of $\Sigma \Rightarrow C$, follow from the above.

## 4. PROPERTIES RELEVANT TO MULTIVARIATE IFRA DISTRIBUTIONS

The class $C$ of MIFRA distributions has been shown by Block and Savits [3] to satisfy the following properties:

(Property 1): Closure under the formation of monotone systems, i.e. if $(T_1, \ldots, T_n) \in C$ and $\tau_1, \ldots, \tau_m$ are monotone life functions, then $(\tau_1(T_1, \ldots, T_n), \ldots, \tau_m(T_1, \ldots, T_n)) \in C$.

(Property 2): Closure under limits in distribution.

(Property 3): Marginals are in the same class.

(Property 4): Closure under conjunction of independent sets of lifetimes, i.e. if $(T_1, \ldots, T_n)$ and $(S_1, \ldots, S_m) \in C$ and are independent, then $(T_1, \ldots, T_n, S_1, \ldots, S_m) \in C$.

(Property 5): Closure under scaling, i.e. if $(T_1, \ldots, T_n) \in C$ and $a_i$, $i = 1, \ldots, n$, are nonnegative constants, then $(a_1 T_1, \ldots, a_n T_n) \in C$.

(Property 6): $C$ is closed under well defined convolution, i.e. if $(T_1, \ldots, T_n) \in C$ and $(S_1, \ldots, S_n) \in C$ and independent, then $(T_1 + S_1, \ldots, T_n + S_n) \in C$.

It is reasonable that any class of multivariate IFRA distributions should satisfy these conditions. Block and Savits [3] have shown that the MIFRA distributions satisfy these conditions. We will now show that each of the conditions $A$, $B$, $C$, $\Sigma$, $D$, $E$, $F$ fails to satisfy at least one of these properties.

### 4.1 $A$ Does Not Satisfy Property 1.

This follows since $A \not\Rightarrow B$.

### 4.2 $B$ Does Not Satisfy Property 5.

This follows since $B \not\Rightarrow A$.

### 4.3 $C$ (and $D$) Do Not Satisfy Property 5.

Let $(T_1, T_2) = (\min(X, Z), \min(Y, Z))$ where $X$, $Y$ and $Z$ are independent exponential random variables with mean one. For $a_1 \neq a_2$, assume $(a_1 T_1, a_2 T_2) = (\tau_1(X_1, \ldots, X_k), \tau_2(X_1, \ldots, X_k))$ where $\tau_1$, $\tau_2$ are monotone life functions and $X_1, \ldots, X_k$ are independent IFRA lifetimes. It follows from Remark 2.2a of Block and Savits [4] that there exist independent exponential random variables $U$, $V$, $W$ such that $(a_1 T_1, a_2 T_2) = (\min(U, W), \min(V, W))$. But the conditions $0 = P(aT_1 = a_2 T_2)$ and $P(\min(U, W) = \min(V, W)) > 0$ are not compatible.

### 4.4 $\Sigma$ Does Not Satisfy Property 1.

Let $X$ and $Y$ be independent exponential lifetimes. Define $\tau_1(X, Y) = \min(X, Y)$ and $\tau_2(X, Y) = Y$ and assume that $(\tau_1, \tau_2) = (U + W, V + W)$ where $U$, $V$, and $W$ are independent IFRA lifetimes. Now by Theorem 2.8 of Block and Savits [4], one of $V$ and $W$ is exponential and one is concentrated at 0. If $W$ is exponential since $P(\min(X, Y) \leqslant Y) = 1$, it follows that $P(U = 0) = 1$ and so $P(\min(X, Y) = Y) = 1$ which is impossible if $X$ and $Y$ are independent exponentials. If $V$ is exponential, then $P(W = 0) = 1$ so that $\min(X, Y)$ and $Y$ are independent, again an impossibility.

### 4.5 $D$ Does Not Satisfy Property 6.

Let $X$, $Y$, and $Z$ be independent absolutely continuous IFRA lifetimes. Then both $(X, Y)$ and $(Z, Z)$ are trivially in $D$. However, if $(X, Y) + (Z, Z) = (X + Z, Y + Z)$ was in $D$, then by Section 10 of Esary and Marshall [6] $X + Z$ and $Y + Z$ would be independent, but they can't be.

### 4.6 $E$ Does Not Satisfy Property 1.

This follows since $E \not\Rightarrow B$.

### 4.7 $F$ Does Not Satisfy Property 1.

This follows since $A \Leftrightarrow F$.

## 5. POSITIVE DEPENDENCE

The first published definition of a class of multivariate nonparametric reliability distributions was Harris' [7] definition for multivariate increasing hazard rate. This definition included a type of positive dependence (i.e. right corner set increasing). See Barlow and Proschan [1] for a discussion of various types of positive dependence. Subsequent definitions have not included such assumptions. The opinion which is now generally held is that the various concepts of positive dependence are not intimately related to useful definitions for nonparametric multivariate life classes. In other words, if a multivariate lifetime has an increasing failure rate or failure rate average, then it need not follow that the lifetime be positively dependent in some sense. In fact, if such a definition implies positive dependence, then it is probably too strong. Examples of such definitions are conditions $C$, $D$ and $\Sigma$ which are easily shown to imply association. We will show that the more useful definitions $A\ (\equiv F)$, $B$, $E$ and especially MIFRA do not imply even positive quadrant dependence, which is one of the weaker types of positive dependence.

### 5.1 $A$ and $E \not\Rightarrow$ Positive Quadrant Dependence

Clearly, $\bar{F}(t_1, t_2) = P(T_1 > t_1, T_2 > t_2) = \exp(-t_1 - t_2 - t_1 t_2)$ satisfies $A$ and $E$, but $\bar{F}(t_1, t_2) \leqslant P(T_1 > t_1)P(T_2 > t_2)$.

### 5.2 MIFRA and $B \not\Rightarrow$ Positive Quadrant Dependence

Consider $(T_1, T_2) = (U, 1 - U)$ where $U$ is a uniform distribution on the unit interval. Clearly, $\bar{F}(t_1, t_2) \leqslant P(T_1 > t_1)P(T_2 > t_2)$, but $T_1$, $T_2$, $\min(T_1, T_2)$ and $\max(T_1, T_2)$ have a univariate uniform distribution and so are IFRA. Thus, $B$ is satisfied. Furthermore, Theorem 3.5 of Block and Savits [3] gives that $(T_1, T_2)$ is MIFRA if the indicator function of every fundamental upper domain in $\mathbb{R}_2^+$ satisfies the inequality of Definition 1. A set $A$ is a fundamental upper domain if it has the form in Figure 2 below where $0 \leqslant x_1 \leqslant x_2 \leqslant \ldots \leqslant x_n$ and $y_1 \geqslant y_2 \geqslant \ldots \geqslant y_n \geqslant 0$. From Figure 2,

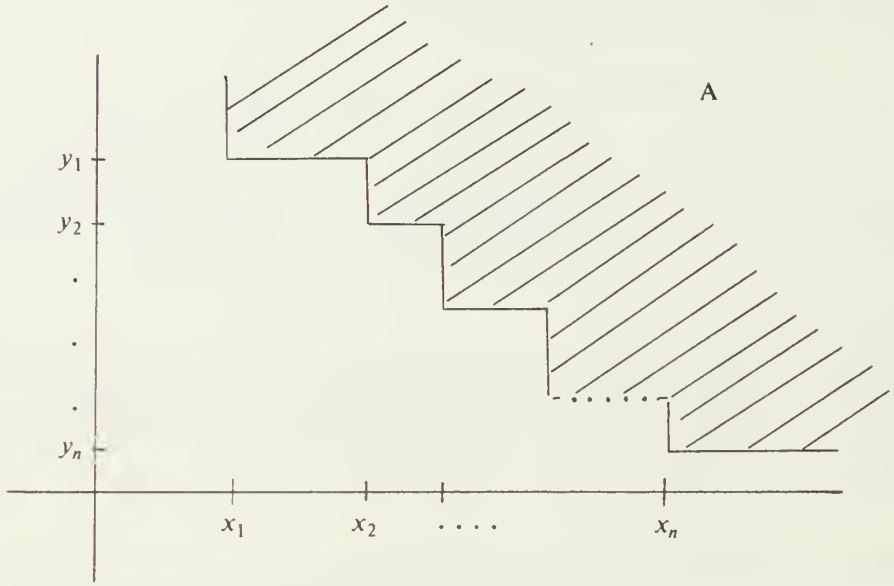



FIGURE 2. Fundamental upper domain

$$\{(T_1,\ T_2) \in A\} = \bigcup_{i=1}^{n} \{T_1 > x_i,\ T_2 > y_i\} = \bigcup_{i=1}^{n} \{x_i < U < 1 - y_i\}$$

and

$$\{(T_1/\alpha,\ T_2/\alpha) \in A\} = \bigcup_{i=1}^{n} \{\alpha x_i < U < 1 - \alpha y_i\}.$$

Let $I = \{i\colon x_i + y_i < 1\}$, $J = \{j\colon \alpha x_j + \alpha y_j < 1\}$. Since $0 < \alpha \leqslant 1$, $I \subseteq J$. Then

$$E[I_A(T_1,\ T_2)] = P\left[\bigcup_{i \in I} \{x_i < U < 1 - y_i\}\right]$$

and

$$\begin{aligned} E^{1/\alpha}[I_A^{\alpha}(T_1/\alpha,\ T_2/\alpha)] &= P^{1/\alpha}\left[\bigcup_{i \in I} \{\alpha x_i < U < 1 - \alpha y_i\}\right] \\ &\geqslant P^{1/\alpha}\left[\bigcup_{i \in I} \{\alpha x_i < U < 1 - \alpha y_i\}\right]. \end{aligned}$$

By renumbering, if necessary, we may assume without loss of generality that $I = \{1, 2, \ldots, p\}$. Now define $K = \{2 \leqslant k \leqslant p\colon \alpha x_k \geqslant 1 - \alpha y_{k-1}\} = \{k_1 < k_2 < \ldots < k_r\}$ and set $k_0 = 1$ and $k_{r+1} = p + 1$. Then

$$\bigcup_{i \in I} \{\alpha x_i < U < 1 - \alpha y_i\} = \bigcup_{l=1}^{r+1} \{\alpha x_{k_{l-1}} < U < 1 - \alpha y_{k_l - 1}\}$$

and these latter sets are disjoint intervals. It follows from Minkowski's inequality for $0 < \alpha \leqslant 1$ that

$$\begin{aligned} P^{1/\alpha}\left[\bigcup_{i \in I} \{\alpha x_i < U < 1 - \alpha y_i\}\right] &= \left\{\sum_{l=1}^{r+1} P\{\alpha x_{k_{l-1}} < U < 1 - \alpha y_{k_l - 1}\}\right\}^{1/\alpha} \\ &\geqslant \sum_{l=1}^{r+1} P^{1/\alpha}\{\alpha x_{k_{l-1}} < U < 1 - \alpha y_{k_l - 1}\}. \end{aligned}$$

Since

$$\bigcup_{i \in I} \{x_i < U < 1 - y_i\} = \bigcup_{l=1}^{r+1} \bigcup_{k_{l-1} \leqslant i \leqslant k_l - 1} \{x_i < U < 1 - y_i\},$$

it suffices to show that for $l = 1, \ldots, r + 1$

$$P\left[\bigcup_{k_{l-1} \leqslant i \leqslant k_l - 1} \{x_i < U < 1 - y_i\}\right] \leqslant P^{1/\alpha} \{\alpha x_{k_{l-1}} < U < 1 - \alpha y_{k_l - 1}\}.$$

But the union on the left hand side is contained in the interval $\{x_{k_{l-1}} < U < 1 - y_{k_l - 1}\}$ and $x_{k_{l-1}} + y_{k_l - 1} < 1$ since for $k_{l-1} \leqslant i \leqslant k_l - 1$, $i \in I$. And so we are done if we show that $1 - t \leqslant (1 - \alpha t)^{1/\alpha}$ for $0 \leqslant t \leqslant 1$, $0 < \alpha \leqslant 1$. This last inequality is, however, easily verified.

## REFERENCES


[1] Barlow, R.E. and F. Proschan, *Statistical Theory of Reliability and Life Testing: Probability Models.* (Holt, Rinehart and Winston, New York, 1975).

[2] Block, H.W. and T.H. Savits, "The IFRA Closure Problem," Annals of Probability, *4*, 1030-1032 (1976).

[3] Block, H.W. and T.H. Savits, "Multivariate IFRA Distributions," Annals of Probability, *8*, 793-801 (1980).

[4] Block, H.W. and T.H. Savits, "Systems with Exponential System Life and IFRA Component Lives," Annals of Statistics, *7*, 911-916 (1979).

[5] Esary, J.D. and A.W. Marshall, "Coherent Life Functions," SIAM Journal of Applied Mathematics, *18*, 810-814 (1970).

[6] Esary, J.D. and A.W. Marshall, "Multivariate Distributions with Increasing Hazard Rate Average," Annals of Probability, *7*, 359-370 (1979).

[7] Harris, R., "A Multivariate Definition for Increasing Hazard Rate Distribution Functions," Annals of Mathematical Statistics, *41*, 713-717 (1970).

[8] Marshall, A.W. and M. Shaked, "A Class of Multivariate New Better than Used Distributions," Annals of Probability, *10*, 259-264 (1982).

# REMARKS ON A UNIVARIATE SHOCK MODEL WITH SOME BIVARIATE GENERALIZATIONS

William S. Griffith

*University of Kentucky*
*Department of Statistics*
*Lexington, Kentucky*

**ABSTRACT**

In a 1973 paper J. D. Esary, A. W. Marshall, and F. Proschan [5] considered a shock model giving rise to various nonparametric classes of life distributions of interest in reliability theory. A number of authors have extended these results in a variety of directions. In this paper, alternative proofs of the increasing failure rate (IFR) and decreasing mean residual life (DMRL) results are given which do not make use of the theory of total positivity. Some bivariate extensions are then obtained using a shock model similar to that originally used by H. W. Block, A. S. Paulson, and R. C. Kohberger [2] to unify various bivariate exponential distributions.

## INTRODUCTION

In a 1973 paper, Esary, Marshall, and Proschan [5] investigated a shock model which gives rise to various nonparametric classes of life distributions which are of interest in reliability theory. We will begin this paper by reviewing these univariate life distribution classes and then stating the theorem due to Esary, Marshall, and Proschan. We then provide new proofs to some parts of this theorem. The proof uses elementary methods rather than results from the theory of total positivity. We will then describe a bivariate shock model which has been used by Block, Paulson, and Kohberger [2] to unify some of the various bivariate exponential distributions which have appeared in the reliability literature. Finally, we will discuss a number of bivariate shock models yielding bivariate conditions that can be viewed as generalizations of the univariate new better than used (NBU) condition.

## 1. UNIVARIATE LIFE DISTRIBUTION CLASSES

In this section, we will consider various models of stochastic aging. For every model of wearout (or adverse stochastic aging) there is an analogous model of improvement (or beneficial stochastic aging). Throughout, we will use the word "increasing" ("decreasing") to mean "nondecreasing" ("nonincreasing").

Consider a single component or system with a random lifetime denoted by the random variable $T$. We will denote the distribution function of $T$ by $F$. The survival function, denoted by $\bar{F}$, is $\bar{F}(t) = P(T > t) = 1 - F(t)$. We will assume throughout that $\bar{F}(0-) = 1$ (i.e., that $T$ is a nonnegative random variable). We now define the various nonparametric life distribution classes of interest.

If a life distribution $F$ satisfies

$$\frac{\bar{F}(t+x)}{\bar{F}(t)} \text{ is decreasing in } -\infty < t < \infty \text{ for fixed } x > 0 \tag{1.1}$$

then $F$ is called an increasing failure rate (IFR) distribution.

A distribution $F$ is called an increasing failure rate average (IFRA) distribution if

$$\bar{F}(\alpha t) \geq \bar{F}^{\alpha}(t) \text{ for } 0 < \alpha < 1, \quad t \geq 0. \tag{1.2}$$

A distribution $F$ is called a new better than used (NBU) distribution if

$$\bar{F}(x + y) \leq \bar{F}(x)\bar{F}(y) \text{ for } x \geq 0, \quad y \geq 0. \tag{1.3}$$

A distribution $F$ with a finite mean $\mu$ is called a new better than used in expectation (NBUE) distribution if

$$\int_t^{\infty} \bar{F}(x)dx \leq \mu \bar{F}(t) \text{ for } t \geq 0. \tag{1.4}$$

A distribution $F$ with finite mean is called a decreasing mean residual life (DMRL) distribution if

$$\frac{1}{\bar{F}(t)} \int_t^{\infty} \bar{F}(x)dx \text{ is decreasing in } t. \tag{1.5}$$

For each of the preceding classes of life distributions, there is a corresponding class which is a model of beneficial stochastic aging. These are obtained from the preceding classes by replacing decreasing with increasing, increasing with decreasing, and reversing inequalities. This results in classes known as the decreasing failure rate (DFR) class, the decreasing failure rate average (DFRA) class, the new worse than used (NWU) class, the new worse than used in expectation (NWUE) class, and the increasing mean residual life (IMRL) class. We next discuss some alternate characterizations which perhaps supply more intuition.

We begin by considering again a random variable $T$ with distribution function $F$ denoting the random lifelength of some device. For any $x \geq 0$, we will consider the survival function

$$\bar{F}_x(t) = \frac{\bar{F}(x + t)}{\bar{F}(x)} \tag{1.7}$$

Intuitively, $\bar{F}_x$ represents the survival function of the residual life of the original device conditioned on the fact that it survived past time $x$. We next prove a lemma and a number of propositions which give alternate characterizations of most of these classes.

LEMMA 1:

$$\bar{F}_y = (\bar{F}_x)_{y-x} \text{ for } y \geq x \geq 0$$

PROOF: Omitted.

PROPOSITION 1: Let $\mu_x$ and $\mu$ be the means of $F_x$ and $F$, respectively.

(i) $F$ has an IFR distribution if and only if $\bar{F}_x(z) \leq \bar{F}_y(z)$ for all $z$ when $x \geq y$.

(ii) $F$ has an NBU distribution if and only if $\bar{F}_x(z) \leq \bar{F}(z)$ for all $z$ when $x \geq 0$.

Let $\mu$ be finite

(iii) $F$ is a DMRL distribution if and only if $\mu_x \leq \mu_y$ for $x \geq y \geq 0$.

(iv) $F$ is an NBUE distribution if and only if $\mu_x \leq \mu$ for $x \geq 0$.

PROOF: Omitted.

PROPOSITION 2:

(i) $F$ is an IFR distribution if and only if $F_y$ is an NBU distribution for all $y$.

(ii) Let $F$ have a finite mean. Then $F$ has a DMRL distribution if and only if $F_y$ has an NBUE distribution for all $y$.

PROOF: The proof of (i) follows directly from Proposition 1(i), Lemma 1, and Proposition 1(ii). The proof of (ii) follows from Proposition 1(iii), Lemma 1, and Proposition 1(iv).

## 2. THE EMP SHOCK MODEL AND THEOREM

In this section, we will discuss the Esary, Marshall, and Proschan (EMP) shock model and theorem. We will provide alternate proofs to some of these results and indicate some generalizations.

We begin by defining various nonparametric classes of discrete distributions for positive integer valued random variables.

Let $M$ be a positive integer valued random variable putting no mass at infinity. Then we shall say that

(2.1) $M$ is a discrete IFR random variable if $(P(M > k)/P(M > k-1)$ is decreasing in $k = 1, 2, \ldots$.

(2.2) $M$ is a discrete IFRA random variable if $\{P(M > k)\}^{1/k}$ is decreasing in $k = 1, 2, \ldots$.

(2.3) $M$ is a discrete NBU random variable if $P(M > k + j) \leqslant P(M > k)P(M > j)$ for $j, k = 0, 1, 2, \ldots$.

(2.4) $M$ is a discrete NBUE random variable if $P(M > k) \sum_{j=0}^{\infty} P(M > j) \geqslant \sum_{j=k}^{\infty} P(M > j)$ for $k = 0, 1, 2, \ldots$ and $M$ has a finite mean.

(2.5) $M$ is a discrete DMRL random variable if $\{P(M > k)\}^{-1} \left\{ \sum_{j=k}^{\infty} P(M > j) \right\}$ is decreasing in $k = 0, 1, 2, \ldots$ and $M$ has a finite mean.

Note that these classes of life distributions are simply discrete analogues of the classes that we have previously defined in Section 1. If we reverse the inequalities and change decreasing to increasing we can define analogous classes of discrete DFR, DFRA, NWU, NWUE, and IMRL distributions (except that we do not require $M$ to have a finite mean in the NWUE and IMRL cases).

Next we describe the fundamental shock model discussed by Esary, Marshall, and Proschan. Consider a device subjected to shocks which occur in time as events of a Poisson process with intensity $\lambda$. The device fails on the $M$th shock where $M$ is a positive integer valued random variable which is independent of the shock process. Letting $T$ be the lifelength of the device, we see that $T = \sum_{i=1}^{M} X_i$ where the $X_i$'s are independent and identically distributed random variables having an exponential distribution with parameter $\lambda$. The survival function is of the form

$$\bar{F}(t) = P(T > t) = \sum_{k=0}^{\infty} P(S(t) = k)P(M > k)$$

where $\{S(t): t \geqslant 0\}$ is a Poisson Process with intensity $\lambda$.

We now state the main theorem.

THEOREM 1: (Esary, Marshall, and Proschan):

(i) If $M$ is discrete IFR (DFR), then $F$ is IFR (DFR)

(ii) If $M$ is discrete IFRA (DFRA), then $F$ is IFRA (DFRA)

(iii) If $M$ is discrete NBU (NWU), then $F$ is NBU (NWU)

(iv) If $M$ is discrete NBUE (NWUE), then $F$ is NBUE (NWUE)

(v) If $M$ is discrete DMRL (IMRL), then $F$ is DMRL (IMRL).

Elementary proofs of parts (iii) and (iv) were given in Esary, Marshall, and Proschan. The proofs of the remaining parts required results from the theory of total positivity. We now give proofs of (i) and (v) which involve elementary methods. We repeat the proof of (iii) to illustrate some techniques which we will utilize. Furthermore, in the subsequent section, we use the techniques developed to establish new bivariate results.

PROOF OF THEOREM 1(iii): We note that

$$\bar{F}(t + x) = \sum_{k=0}^{\infty} P(S(t + x) = k)\, P(M > k).$$

Since $\{S(t): t \geqslant 0\}$ is assumed to be a Poisson process, we have by the stationary independent increments property that

$$\begin{aligned}\bar{F}(t + x) &= \sum_{k=0}^{\infty} \sum_{\Delta=0}^{k} P(S(t) = \Delta)P(S(x) = k - \Delta)P(M > \Delta + (k - \Delta)) \\ &= \sum_{\Delta=0}^{\infty} \sum_{k=\Delta}^{\infty} P(S(t) = \Delta)P(S(x) = k - \Delta)P(M > \Delta + (k - \Delta)) \\ &= \sum_{\Delta=0}^{\infty} \sum_{l=0}^{\infty} (P(S(t) = \Delta)P(S(x) = l)P(M > \Delta + l).\end{aligned}$$

Applying the discrete NBU condition on $M$ we have that

$$\begin{aligned}\bar{F}(t + x) &\leqslant \sum_{\Delta=0}^{\infty} \sum_{l=0}^{\infty} P(S(t) = \Delta)P(S(x) = l)P(M > \Delta)P(M > l) \\ &= \left[\sum_{\Delta=0}^{\infty} P(S(t) = \Delta)P(M > \Delta)\right]\left[\sum_{l=0}^{\infty} P(S(x) = l)P(M > l)\right].\end{aligned}$$

Consequently, $\bar{F}(t + x) \leqslant \bar{F}(t)\bar{F}(x)$.

PROOF OF THEOREM 1 (i): Consider a random variable $T$ whose life distribution arises from a shock model in which $M$ is discrete IFR. In order to show that it is IFR It suffices from Proposition 2 that $F_y$ is NBU for all $y$. However, in light of the preceding proof, we need only show that $F_y$ arises from a shock model in which the random number of shocks survived has a discrete NBU distribution.

Now

$$\bar{F}_y(u) = \frac{\bar{F}(y+u)}{\bar{F}(y)} = \frac{\sum_{\Delta=0}^{\infty} P(S(y+u)=\Delta)P(M>\Delta)}{\sum_{j=0}^{\infty} P(S(y)=j)P(M>j)}$$

$$= \frac{\sum_{\Delta=0}^{\infty}\sum_{j=0}^{\Delta} P(S(y)=j,\ S_y(u)=\Delta-j)P(M>j+(\Delta-j))}{\sum_{j=0}^{\infty} P(S(y)=j)P(M>j)}$$

where $S_y(u) = S(y+u) - S(y)$

$$= \frac{\sum_{j=0}^{\infty}\sum_{\Delta=j}^{\infty} P(S(y)=j)P(S_y(U)=\Delta-j)P(M>j+(\Delta-j))}{\sum_{j=0}^{\infty} P(S(y)=j)P(M>j)}$$

$$= \frac{\sum_{j=0}^{\infty}\sum_{k=0}^{\infty} P(S(y)=j)P(S_y(u)=k)P(M>j+k)}{\sum_{j=0}^{\infty} P(S(y)=j)P(M>j)}$$

$$= \sum_{k=0}^{\infty} P(S_y(u)=k)\,\frac{\sum_{j=0}^{\infty} P(S(y)=j)P(M>j+k)}{\sum_{j=0}^{\infty} P(S(y)=j)P(M>j)}$$

Now the expression

$$\frac{\sum_{j=0}^{\infty} P(S(y)=j)P(M>j+k)}{\sum_{j=0}^{\infty} P(S(y)=j)\,(P(M>j)} \tag{2.6}$$

is a function of $k$ which has the value 1 when $k=0$ and is a decreasing function of $k$ since $P(M>j+k)$ decreases in $k$ for fixed $j$. In addition, since $M$ is discrete IFR, it is also discrete NBU and hence $P(M>j+k) \leqslant P(M>j)P(M>k)$. Thus, the expression in equation 2.6 is bounded between 0 and $P(M>k)$. However, as $k \to \infty$, $P(M>k) \to 0$. Consequently, the expression in equation 2.6 is a survival function of a discrete random variable.

Next we let $M^*$ be a discrete random variable which is independent of the Poisson process $\{S_y(t): t \geqslant 0\}$ which has this survival function. Then $\bar{F}_y(u)$ is of the form

$$\bar{F}_y(u) = \sum_{k=0}^{\infty} P(S_y(u)=k)P(M^*>k).$$

To complete the proof we need only show that since $M$ is discrete IFR, $M^*$ must be discrete NBU.

The condition that must be checked is

$$P(M^*>k+l) \leqslant P(M^*>k)P(M^*>l).$$

This is equivalent to

$$\frac{\sum_{j=0}^{\infty} P(S(y)=j)P(M>j+k+l)}{\sum_{j=0}^{\infty} P(S(y)=j)P(M>j)} \leqslant \frac{\sum_{j=0}^{\infty} P(S(y)=j)P(M>j+k)}{\sum_{j=0}^{\infty} P(S(y)=j)P(M>j)} \frac{\sum_{j=0}^{\infty} P(S(y)=j)P(M>j+l)}{\sum_{j=0}^{\infty} P(S(y)=j)P(M>j)}$$

or after some algebra to

$$\sum_{\Delta=0}^{\infty}\sum_{j=0}^{\infty}(P(M>j+k+l)P(M>\Delta)-P(M>j+k)P(M>\Delta+l))P(S(y)=\Delta)P(S(y)=j)\leqslant 0.$$

Since for $\Delta < 0$ or $j < 0$, $P(S(y)=\Delta)P(S(y)=j)=0$, we may rewrite the above condition as

$$\sum_{\Delta=-\infty}^{\infty}\sum_{j=-\infty}^{\infty}(P(M>j+k+l)P(M>\Delta)-P(M>j+k)P(M>\Delta+l))P(S(y)=\Delta)P(S(y)=j)\leqslant 0.$$

By breaking up this double summation into three parts: over $\Delta < j+k$, $\Delta = j+k$, and $\Delta > j+k$, realizing tht the middle term of the preceding sum is zero; letting $\Delta' = j+k$ and $j' = \Delta - k$ so that $j = \Delta' - k$ and $\Delta = j' + k$ (with the consequence that the condition $\Delta > j+k$ is equivalent with $j' + k > \Delta'$) and rearranging, we then find that the condition we are checking can be rewritten as

$$\sum_{\Delta<j+k}\sum(P(M>j+k+l)P(M>\Delta)-P(M>j+k)P(M>\Delta+l))P(S(y)=\Delta)P(S(y)=j)$$
$$-\sum_{\Delta<j+k}\sum\{(P(M>j+k+l)P(M>\Delta)-P(M>j+k)P(M>\Delta+l))$$
$$P(S(y)=j+k)P(S(y)=\Delta-k)\}\leqslant 0,$$

or

$$\sum_{\Delta<j+k}\sum\{(P(M>j+k+l)P(M>\Delta)-P(M>j+k)P(M>\Delta+l))$$
$$(P(S(y)=\Delta)P(S(y)=j)-P(S(y)=j+k)P(S(y)=\Delta-k))\}\leqslant 0.$$

Noticing that

$$P(M>j+k+l)P(M>\Delta)-P(M>j+k)P(M>\Delta+l)$$
$$=\left(\frac{P(M>j+k+l)}{(P(M>j+k)}-\frac{P(M>\Delta+l)}{P(M>\Delta)}\right)P(M>j+k)P(M>\Delta)\leqslant 0$$

since $M$ is discrete IFR and $\Delta < j+k$, we see that it now suffices to show only that

$$P(S(y)=\Delta)P(S(y)=j)-P(S(y)=j+k)P(S(y)=\Delta-k)\geqslant 0. \tag{2.7}$$

We may take care of a number of cases rather easily. If $\Delta < 0$, then $\Delta - k < 0$ and the left hand side of (2.7) is $0-0=0$. If $\Delta \geqslant 0$ and $j < 0$, then $\Delta - k < 0$ since $\Delta < j+k$. Thus, the left hand side of (2.7) is $0-0=0$. If $0 \leqslant \Delta < k$ and $j \geqslant 0$, then the expression on the left hand side of (2.7) reduces to the nonegative quantity $P(S(y)=\Delta)P(S(y)=j)$. Finally, we consider the case where $\Delta \geqslant k$ and $j \geqslant 0$. First, note that if $\Delta \geqslant k$ and $\Delta < j+k$, then $j \neq 0$. If $k=0$ then (2.7) is immediate, thus we consider $k > 0$. The left hand side of (2.7) is

$$\frac{e^{-\lambda y}(\lambda y)^{\Delta}}{\Delta!}\frac{e^{-\lambda y}(\lambda y)^{j}}{j!}-\frac{e^{-\lambda y}(\lambda y)^{j+k}}{(j+k)!}\frac{e^{-\lambda y}(\lambda y)^{\Delta-k}}{(\Delta-k)!}$$
$$=\frac{e^{-2\lambda y}(\lambda y)^{\Delta+j}}{\Delta!\,(j+k)!}\{(j+k)(j+k-1)\ldots(j+1)-\Delta(\Delta-1)\ldots(\Delta-k+1)\}$$

which is clearly nonnegative if $\Delta < j+k$. This completes the proof.

Next, for the proof of Theorem 1 (iv), we refer the reader to the paper of Esary, Marshall, and Proschan. We now show that (v) of Theorem 1 follows from (iv) in much the same fashion that (i) followed from (iii).

PROOF OF THEOREM 1 (v): Consider a random variable $T$ whose life distribution arises from a shock model where $M$ is discrete DMRL. In order to show that it is DMRL, it suffices from Proposition 2 to show that $F_y$ is an NBUE distribution for all $y$. However, in light of (iv), we need only show that $F_y$ arises from a shock model in which the number of shocks survived has a discrete NBUE distribution. Now as before it can be shown that

$$\bar{F}_y(u) = \sum_{k=0}^{\infty} \left\{ P(S_y(u) = k) \frac{\sum_{j=0}^{\infty} P(S(y) = j)P(M > j + k)}{\sum_{j=0}^{\infty} P(S(y) = j)P(M > j)} \right\}.$$

Again, we can show that the expression

$$\frac{\sum_{j=0}^{\infty} P(S(y) = j)P(M > j + k)}{\sum_{j=0}^{\infty} P(S(y) = j)P(M > j)}$$

is the survival function for some discrete distribution.

Now if, as before, we let $M^*$ be a discrete random variable which is independent of the Poisson process $\{S_y(t): t \geq 0\}$ and has this survival function, then $\bar{F}_y(u)$ is of the form

$$\bar{F}_y(u) = \sum_{k=0}^{\infty} P(S_y(u) = k)P(M^* > k).$$

To complete the proof it will suffice to show that $M^*$ is discrete NBUE since $M$ is discrete DMRL. That is, it will suffice to show that

$$\{P(M^* > l)\}^{-1} \sum_{k=l}^{\infty} P(M^* > k) \leq \sum_{k=0}^{\infty} P(M^* > k). \tag{2.8}$$

This can be simplified to

$$\frac{\sum_{j=0}^{\infty} P(S(y) = j) \sum_{k=l}^{\infty} P(M > j + k)}{\sum_{j=0}^{\infty} P(S(y) = j)P(M > j + l)} \leq \frac{\sum_{j=0}^{\infty} \sum_{k=0}^{\infty} P(S(y) = j)P(M > j + k)}{\sum_{j=0}^{\infty} P(S(y) = j)P(M > j)}. \tag{2.9}$$

Now after some algebra we may rewrite the condition from (2.9) as

$$\sum_{j=0}^{\infty} \sum_{\Delta=0}^{\infty} \left[ P(S(y) = j)P(S(y) = \Delta)P(M > j + l)P(M > \Delta) \right.$$
$$\left. \left\{ \frac{1}{P(M > j + l)} \sum_{k=j+l}^{\infty} P(M > k) - \frac{1}{P(M > \Delta)} \sum_{k=\Delta}^{\infty} P(M > k) \right\} \right] \leq 0.$$

Since for $j > 0$ or $\Delta < 0$, $P(S(y) = j)P(S(y) = \Delta) = 0$, we may rewrite the preceding as

$$\sum_{j=-\infty}^{\infty} \sum_{\Delta=-\infty}^{\infty} \left[P(S(y) = j)P(S(y) = \Delta)P(M > j + l)P(M > \Delta) \right.$$
$$\left. \left\{ \frac{1}{P(M > j + l)} \sum_{k=j+l}^{\infty} P(M > k) - \frac{1}{P(M > \Delta)} \sum_{k=\Delta}^{\infty} P(M > k) \right\} \right] \leqslant 0.$$

By breaking the summation into three parts and using methods analogous to the IFR case we can justify this inequality and thus complete the proof.

## 3. A BIVARIATE SHOCK MODEL

Now consider two devices which are subject to shocks. We will denote the number of shocks to devices 1 and 2 in $[0, t]$ by $S_1(t)$ and $S_2(t)$, respectively. Further, we let $(M_1, M_2)$ denote the random number of shocks until failure of devices 1 and 2, respectively. Block, Paulson, and Kohberger [2] have unified some of the various bivariate exponential distributions in the reliability literature by taking $(M_1, M_2)$ to be a bivariate geometric distribution with joint survival function given by

$$P(M_1 > m_1, M_2 > m_2) = \begin{cases} p_{11}^{m_1} (p_{01} + p_{11})^{m_2 - m_1} & \text{if } m_1 \leqslant m_2 \\ p_{11}^{m_2} (p_{10} + p_{11})^{m_1 - m_2} & \text{if } m_2 \leqslant m_1 \end{cases}$$

where $p_{00} + p_{01} + p_{10} + p_{11} = 1$, $p_{10} + p_{11} < 1$, and $p_{01} + p_{11} < 1$. If we take $S_1(t) \equiv S_2(t)$ to be the same Poisson process, then $(T_1, T_2)$ will have a Marshall and Olkin bivariate exponential distribution. If $S_1(t)$ and $S_2(t)$ are taken to be independent Poisson processes the resulting joint distribution of $(T_1, T_2)$ turns out to be of the type which has been discussed by Downton [4], Hawkes [6], and Paulson [8] depending on the choices of the various parameters. Consequently, in this section, the model in which the shocks occur simultaneously to the two devices, i.e., $S_1(t) \equiv S_2(t)$, will be called the Marshall-Olkin type of shock model, while the model in which the shocks occur according to independent Poisson processes will be called the Downton-Hawkes-Paulson type of shock model. We will now consider various bivariate discrete NBU conditions on $(M_1, M_2)$ which yield bivariate NBU conditions of the type considered by Buchanan and Singpurwalla [3].

### Marshall-Olkin Shock Models

Consider two devices subject to simultaneous shocks occurring according to a Poisson process. Letting $M_1$ and $M_2$ represent the number of shocks until failure of devices 1 and 2, respectively, we will prove some results about $H(t_1, t_2) = P(T_1 > t_1, T_2 > t_2)$ where $T_1$ and $T_2$ are the respective lifetimes of the devices.

PROPOSITION 3: If $P(M_1 > m_1 + j, M_2 > m_2 + j) \leqslant P(M_1 > m_1, M_2 \geqslant m_2)P(M_1 > j, M_2 > j)$ for all nonnegative integers $m_1, m_2, j$, then $\bar{H}(t_1 + x, t_2 + x) \leqslant \bar{H}(t_1, t_2)\bar{H}(x, x)$ for all $t_1, t_2, x \geqslant 0$.

PROOF: We will prove this for $t_1 \leqslant t_2$. The proof for the case where $t_1 \geqslant t_2$ is analogous. Throughout the proof we will make crucial use of the stationary independent increments of the Poisson process. We begin by writing for $t_1 \leqslant t_2$

$$\bar{H}(t_1, t_2) = \sum_{k_1=0}^{\infty} \sum_{k_2=k_1}^{\infty} P(S(t_1) = k_1, S(t_2) = k_2)P(M_1 > k_1, M_2 > k_2).$$

By using the stationary independent increments property, we may rewrite this as

$$\bar{H}(t_1, t_2) = \sum_{k=0}^{\infty} P(S(t_1) = k) \sum_{i=0}^{\infty} P(S(t_2 - t_1) = i) P(M_1 > k, M_2 > k + i). \tag{3.1}$$

In an analogous fashion we may then also write for fixed, but arbitrary $x \geq 0$,

$$\bar{H}(t_1 + x, t_2 + x) = \sum_{k=0}^{\infty} \sum_{\Delta=0}^{k} \sum_{i=0}^{\infty} \{P(S(t_1) = \Delta) P(S(x) = k - \Delta) P(S(t_2 - t_1) = i) P(M_1 > k, M_2 > k + i)\}$$

by using the stationary independent increments property. Next by interchanging the orders of summation and performing some algebra it can be shown that

$$\bar{H}(t_1 + x, t_2 + x) = \sum_{\Delta=0}^{\infty} \left\{ P(S(t_1) = \Delta) \sum_{i=0}^{\infty} P(S(t_2 - t_1) = i) P(M_1 > \Delta, M_2 > \Delta + i) \sum_{j=0}^{\infty} P(S(x) = j) \frac{P(M_1 > \Delta + j, M_2 > \Delta + j + i)}{P(M_1 > \Delta, M_2 > \Delta + i)} \right\}.$$

By using the NBU assumption on $(M_1, M_2)$, we have

$$\bar{H}(t_1 + x, t_2 + x) \leq \left\{ \sum_{\Delta=0}^{\infty} P(S(t_1) = \Delta) \sum_{i=0}^{\infty} P(S(t_2 - t_1) = i) P(M_1 > \Delta, M_2 > \Delta + i) \right\} \cdot \left\{ \sum_{j=0}^{\infty} P(S(x) = j) P(M_1 > j, M_2 > j) \right\}.$$

Rewriting the right-hand side of the above inequality we have

$$\bar{H}(t_1 + x, t_2 + x) \leq \bar{H}(t_1, t_2) \bar{H}(x, x). \tag{3.2}$$

We remark in passing that if the direction of the inequality in the hypothesis of the preceding proposition is reversed, then the inequality in the conclusion is reversed. This would be a model of stochastic improvement with age. We also note that the boundary distributions (ones in which equality holds in equation (3.1)) in this case are the distributions with the bivariate loss of memory property. If exponential marginals were required, the boundary would be the Marshall-Olkin bivariate exponential distribution.

PROPOSITION 4: If $P(M_1 > m_1 + j_1, M_2 > m_2 + j_2) \leq P(M_1 > m_1, M_2 > m_2) P(M_1 > j_1, M_2 > j_2)$ for all nonnegative integers $m_1, m_2, j_1, j_2$ with $(m_1 - m_2)(j_1 - j_2) \geq 0$, then $\bar{H}(t_1 + x_1, t_2 + x_2) \leq \bar{H}(t_1, t_2) \bar{H}(x_1, x_2)$ for all $t_1, t_2, x_1, x_2 \geq 0$ with $(t_1 - t_2)(x_1 - x_2) \geq 0$.

PROOF: We first assume that $t_1 \leq t_2$ and $x_1 \leq x_2$. As we have seen in equation (3.1) of the proof of Proposition 3

$$\bar{H}(t_1, t_2) = \sum_{k=0}^{\infty} P(S(t_1) = k) \sum_{i=0}^{\infty} P(S(t_2 - t_1) = i) P(M_1 > k, M_2 > k + i).$$

Now we may write $\bar{H}(t_1 + x_1, t_2 + x_2)$ as

$$\bar{H}(t_1 + x_1, t_2 + x_2) = \sum_{k=0}^{\infty} \sum_{i=0}^{\infty} \{P(S(t_1 + x_1) = k) P(S(t_2 - t_1 + x_2 - x_1) = i) P(M_1 > k, M_2 > k + i)\}.$$

Now by using the stationary independent increments property, interchanging the order of summation, and performing some algebra, we can show that

$$\bar{H}(t_1 + x_1, t_2 + x_2) = \sum_{\Delta=0}^{\infty} \left\{ P(S(t_1) = \Delta) \sum_{n=0}^{\infty} P(S(t_2 - t_1) = n) P(M_1 > \Delta, M_2 > \Delta + n) \right.$$
$$\sum_{j=0}^{\infty} P(S(x_1) = j) \sum_{m=0}^{\infty} \left[ P(S(x_2 - x_1) = m) \right.$$
$$\left. \left. \frac{P(M_1 > \Delta + j, M_2 > \Delta + j + n + m)}{P(M_1 > \Delta, M_2 > \Delta + n)} \right] \right\}.$$

Using the condition on $(M_1, M_2)$ given in the hypothesis of this proposition we then have that

$$\bar{H}(t_1 + x_1, t_2 + x_2) \leqslant \left\{ \sum_{\Delta=0}^{\infty} P(S(t_1) = \Delta) \sum_{n=0}^{\infty} P(S(t_2 - t_1) = n) P(M_1 > \Delta, M_2 > \Delta + n) \right\}.$$
$$\left\{ \sum_{j=0}^{\infty} P(S(x_1) = j) \sum_{m=0}^{\infty} P(S(x_2 - x_1) = m) P(M_1 > j, M_2 > j + m) \right\}.$$

But the right-hand side of the preceding inequality is just $\bar{H}(t_1, t_2)\bar{H}(x_1, x_2)$. Thus, we have for $0 \leqslant t_1 \leqslant t_2$ and $0 \leqslant x_1 \leqslant x_2$ that

$$\bar{H}(t_1 + x_1, t_2 + x_2) \leqslant \bar{H}(t_1, t_2)\bar{H}(x_1, x_2). \tag{3.3}$$

For $t_1 \geqslant t_2 \geqslant 0$ and $x_1 \geqslant x_2 \geqslant 0$, an analogous argument will lead to equation (3.3). The boundary distribution (in which equality holds in equation (3.3)) turns out to be the Marshall-Olkin bivariate exponential distribution. Also a reversal of the inequality in the hypothesis of this proposition leads to a reversal of the inequality in the conclusion. This is a model of stochastic improvement with age.

Marshall and Shaked [7] have considered a cumulative damage model in which the damages $(D_{1i}, D_{2i})$ are independent and identically distributed random vectors with a joint distribution function $G$. Then if $d_1$ and $d_2$ are the fixed failure thresholds of the devices, then the survival function of $(M_1, M_2)$ has the form

$$P(M_1 > m_1, M_2 > m_2 = P\left\{ \sum_{i=1}^{m_1} D_{1i} \leqslant d_1, \sum_{i=1}^{m_2} D_{2i} \leqslant d_2 \right\}.$$

Marshall and Shaked prove that when the survival function of $(M_1, M_2)$ is of this form that the hypothesis of the preceding proposition is satisfied. Consequently, the joint survival function of the devices satisfies equation (3.3) for $(t_1 - t_2)(x_1 - x_2) \geqslant 0$.

PROPOSITION 5: If $P(M_1 > m + j, M_2 > m + j) \leqslant P(M_1 > m, M_2 > m)\ P(M_1 > j, M_2 > j)$ for all nonnegative integers $m$ and $j$, then $\bar{H}(t + x, t + x) \leqslant \bar{H}(t, t)\bar{H}(x, x)$ for all $t, x \geqslant 0$.

PROOF: By the stationary independent increments property we can write

$$\bar{H}(t + x, t + x) = \sum_{k=0}^{\infty} \sum_{j=0}^{\infty} P(S(t) = k) P(S(x) = j) P(M_1 > k + j, M_2 > k + j).$$

Using the hypothesis of the proposition, we have

$$\bar{H}(t + x, t + x) \leqslant \sum_{k=0}^{\infty} \sum_{j=0}^{\infty} P(S(t) = k) P(S(x) = j) P(M_1 > k, M_2 > k) P(M_1 > j, M_2 > j).$$

The right-hand side of the preceding inequality can be rewritten to yield

$$\bar{H}(t + x, t + x) \leqslant \bar{H}(t, t)\bar{H}(x, x).$$

Finally, we will conclude this section with a counterexample. The question is whether a discrete NBU condition

$$P(M_1 > m_1 + j_1, M_2 > m_2 + j_2) \leqslant P(M_1 > m_1, M_2 > m_2)P(M_1 > j_1, M_2 > j_2)$$

for all nonnegative integers $m_1, m_2, j_1, j_2$ would imply the continuous analog

$$\bar{H}(t_1 + x_1, t_2 + x_2) \leqslant \bar{H}(t_1, t_2)\bar{H}(x_1, x_2) \text{ for all } t_1, t_2, x_1, x_2 \geqslant 0.$$

The following counterexample answers this question in the negative. Suppose that $M_1$ and $M_2$ are degenerate random variables with $M_1 = 1$ with probability 1 and $M_2 = 2$ with probability 1. It can be shown that $P(M_1 > m_1 + j_1, M_2 > m_2 + j_2) \leqslant P(M_1 > m_1, M_2 > m_2)P(M_1 > j_1, M_2 > j_2)$ and that if $t_1 < t_2, x_1 > x_2$ and $t_1 + x_1 > t_2 + x_2$,

$$\bar{H}(t_1 + x_1, t_2 + x_2) - \bar{H}(t_1, t_2)\bar{H}(x_1, x_2) = e^{-\lambda(x_1 + t_2)} \sum_{k=2}^{\infty} \frac{(\lambda(t_2 - t_1))^k}{k!}$$

$$> 0.$$

However, an NBU result along these lines does hold in the case of the Downton-Hawkes-Paulson Shock Model.

**A Downton-Hawkes-Paulson Shock Model**

In this model we consider two devices subject to shocks occurring according to independent Poisson processes. Let $\{(S_1(t): t \geqslant 0\}$ and $\{S_2(t): t \geqslant 0\}$ be these processes and $(M_1, M_2)$ be the random number of shocks until failure of the respective devices.

PROPOSITION 6: If $P(M_1 > m_1 + j_1, M_2 > m_2 + j_2) \leqslant P(M_1 > m_1, M_2 > m_2)P(M_1 > j_1, M_2 > j_2)$ for all nonnegative integers $m_1, m_2, j_1, j_2$, then $\bar{H}(t_1 + x_1, t_2 + x_2) \leqslant \bar{H}(t_1, t_2)\bar{H}(x_1, x_2)$ for all $t_1, t_2, x_1, x_2 \geqslant 0$.

PROOF: By the stationary independent increments property and an interchange of the order of summation we have

$$\bar{H}(t_1 + x_1, t_2 + x_2) = \sum_{m=0}^{\infty}\sum_{n=0}^{\infty}\sum_{k=m}^{\infty}\sum_{j=n}^{\infty}\Bigg\{P(S_1(x_1) = m, S_2(x_2) = n)P(M_1 > m, M_2 > n)$$

$$P(S_1(t_1) = k - m, S_2(t_2) = j - n)\frac{P(M_1 > k, M_2 > j)}{P(M_1 > m, M_2 > n)}\Bigg\}.$$

Using the hypothesis of the proposition, we can write

$$\bar{H}(t_1 + x_1, t_2 + x_2) \leqslant \sum_{m=0}^{\infty}\sum_{n=0}^{\infty}\{P(S_1(x_1) = m, S_2(x_2) = n)P(M_1 > m, M_2 > m)$$

$$\sum_{k=m}^{\infty}\sum_{j=n}^{\infty}P(S_1(t_1) = k - m, S_2(t_2) = j - n)$$

$$P(M_1 > k - m, M_2 > j - n)\}.$$

By a change of variables

$$\bar{H}(t_1 + x_1, t_2 + x_2) \leqslant \Bigg\{\sum_{m=0}^{\infty}\sum_{n=0}^{\infty}P(S_1(x_1) = m, S_2(x_2) = n)P(M_1 > m, M_2 > n)\Bigg\}.$$

$$\Bigg\{\sum_{k=0}^{\infty}\sum_{j=0}^{\infty}P(S_1(t_1) = k, S_2(t_2) = j)P(M_1 > k, M_2 > j)\Bigg\}.$$

But the right-hand side of the preceding inequality may be rewritten as $\bar{H}(t_1, t_2)\bar{H}(x_1, x_2)$, thus completing the proof.

## REFERENCES

[1] Barlow, R.E. and F. Proschan, *Statistical Theory of Reliability and Life Testing* (Holt, Rinehart and Winston, New York, 1975).

[2] Block, H.W., A.S. Paulson and R.C. Kohberger, "A Class of Bivariate Distributions" (submitted for publication).

[3] Buchanan, W.B. and N.D. Singpurwalla, "Some Stochastic Characterizations of Multivariate Survival," *The Theory and Applications of Reliability with Emphasis on Bayesian and Nonparametric Methods*," Vol. 1, C. P. Tsokos and I.N. Shimi, Editors (Academic Press, New York, 1977).

[4] Downton, F., "Bivariate Exponential Distributions in Reliability Theory," Journal of the Royal Statistical Society, Series B, *32*, 822-830 (1970).

[5] Esary, J.D., A.W. Marshall and F. Proschan, "Shock Models and Wear Processes," The Annals of Probability, *1*, 627-649 (1973).

[6] Hawkes, A.G., "A Bivariate Exponential Distribution with Applications to Reliability," Journal of the Royal Statistical Society, Series B, *34*, 129-131 (1972).

[7] Marshall, A.W. and M. Shaked, "Multivariate Shock Models for Distributions with Increasing Hazard Rate," The Annals of Probability, *7*, 343-358 (1979).

[8] Paulson, A.S., "A Characterization of the Exponential Distribution and a Bivariate Exponential Distribution," Sankhyā, Series A, *35*, 69-78 (1973).

# ASYMPTOTIC JOINT NORMALITY OF AN INCREASING NUMBER OF MULTIVARIATE ORDER STATISTICS AND ASSOCIATED CELL FREQUENCIES

Lionel Weiss
*Cornell University*
*Ithaca, New York*

**ABSTRACT**

Take $n$ independent identically distributed (IID) observations from a continuous $r$-variate population, and choose some order statistics from each of the $r$ variates. These order statistics are used to construct a grid in $r$-dimensional space. Under certain conditions, it is shown that as $n$ increases we can choose an increasing number of order statistics in such a way that the asymptotic joint distribution of the chosen order statistics and of the frequencies of sample points falling in the cells of the grid can be assumed to be a normal distribution. An application to testing independence of random variables is given.

## 1. NOTATION, ASSUMPTIONS, AND STATEMENT OF THEOREM

We observe $n$ IID $r$-tuplets $(X_1(i), \ldots, X_r(i))$, $i = 1, \ldots, n$, each with joint cumulative distribution function (cdf) $F_n(x_1, \ldots, x_r)$, joint probability density function (pdf) $f_n(x_1, \ldots, x_r)$. We choose a set of order statistics from $(X_1(1), \ldots, X_1(n))$, a set of order statistics from $(X_2(1), \ldots, X_2(n))$, $\ldots$, and a set of order statistics from $(X_r(1), \ldots, X_r(n))$. These chosen order statistics are used to construct a grid in $r$-dimensional space. This grid defines a system of $r$-dimensional cells, each cell having a frequency of sample points falling in its interior. It is clear from the argument of [1] that if the number of chosen order statistics for each variate is fixed, then under mild conditions the joint asymptotic distribution of the chosen order statistics and the cell frequencies (all properly standardized) is normal. But for certain purposes, it is necessary to let the number of order statistics chosen increase as $n$ increases. In this paper we show that under certain conditions we still get a joint asymptotic normal distribution.

For typographical simplicity, until Section 3, we discuss only the case $r = 2$, and we write the $n$ IID pairs observed as $(X_1', Y_1')$, $(X_2', Y_2')$, $\ldots$, $(X_n', Y_n')$. The joint cdf and pdf for $(X_1', Y_1')$ will be written as $F_n(x, y)$, $f_n(x, y)$, respectively, and so are allowed to depend on $n$. $G_n(x)$, $g_n(x)$ denote, respectively, the marginal cdf and pdf for $X_1'$. $H_n(x)$, $h_n(x)$ denote, respectively, the marginal cdf and pdf for $Y_1'$. Define $F_n^{(1)}(x^*, y^*)$ as $\frac{\partial}{\partial x} F_n(x, y)]_{x=x^*, y=y^*}$, and $F_n^{(2)}(x^*, y^*)$ as $\frac{\partial}{\partial y} F_n(x, y)]_{x=x^*, y=y^*}$.

For each $n$, choose $p_n$, $q_n$, and $L_n$ so that $0 < p_n < q_n < 1$, and $np_n$, $nq_n$, $L_n$, and $K_n \equiv \frac{nq_n - np_n}{L_n}$ are all positive integers. These quantities also are to satisfy other conditions to be specified below.

$X_{[1]}' < \ldots < X_{[n]}'$ denote the ordered values of $(X_1', \ldots, X_n')$, and $Y_{[1]}' < \ldots < Y_{[n]}'$ denote the ordered values of $(Y_1', \ldots, Y_n')$. $X_i^*$ denotes $X_{[np_n + (i-1)L_n]}'$ and $Y_i^*$ denotes $Y_{[np_n + (i-1)L_n]}'$ for

$i = 1, \ldots, K_n + 1$. Let $ND(n)$ (for "No duplication in $n$ observations") denote the following event: there is no pair $(X_i', Y_i')$ with $X_i'$ among $(X_1^*, \ldots, X_{K_n+1}^*)$ and $Y_i'$ among $(Y_1^*, \ldots, Y_{K_n+1}^*)$. We assume that $\{f_n, p_n, q_n, L_n\}$ are such that $\lim_{n\to\infty} P[ND(n)] = 1$. This assumption will be discussed in Section 3: under it, for all asymptotic probability calculations, we can (and will) assume that $ND(n)$ occurs for all $n$.

Define $X_0^*$ as $\sup\{x: G_n(x) = 0\}$, $X_{K_n+2}^*$ as $\inf\{x: G_n(x) = 1\}$, $Y_0^*$ as $\sup\{y: H_n(y) = 0\}$, and $Y_{K_n+2}^*$ as $\inf\{y: H_n(y) = 1\}$. Thus, these values may depend on $n$, and may be $-\infty$ or $\infty$.

Define $N_{ij}$ as the number of pairs $(X_i', Y_i')$ in the open rectangle in $(x, y)$-space defined by $(X_{i-1}^* < x < X_i^*,\ Y_{j-1}^* < y < Y_j^*)$, for $i, j$ ranging between 1 and $K_n + 2$, inclusive. Define $S_{ij}$ as 1 if the $Y'$ originally associated with $X_i^*$ falls in the open interval $(Y_{j-1}^*, Y_j^*)$ and as zero otherwise, for $i = 1, \ldots, K_n + 1$ and $j = 1, \ldots, K_n + 2$. Define $T_{ji}$ as 1 if the $X'$ originally associated with $Y_j^*$ falls in the open interval $(X_{i-1}^*, X_i^*)$ and as zero otherwise, for $j = 1, \ldots, K_n + 1$ and $i = 1, \ldots, K_n + 2$. The occurrence of $ND(n)$ then implies all the following equalities:

$$
\begin{aligned}
&\sum_{j=1}^{K_n+2} S_{ij} = 1 \text{ for } i = 1, \ldots, K_n + 1,\\
&\sum_{i=1}^{K_n+2} T_{ji} = 1 \text{ for } j = 1, \ldots, K_n + 1,\\
&\sum_{j=1}^{K_n+2} N_{1j} + \sum_{j=1}^{K_n+1} T_{j1} = np_n - 1,\\
&\sum_{j=1}^{K_n+2} N_{ij} + \sum_{j=1}^{K_n+1} T_{ji} = L_n - 1 \text{ for } i = 2, \ldots, K_n + 1, \qquad (1.1)\\
&\sum_{j=1}^{K_n+2} N_{K_n+2,j} + \sum_{j=1}^{K_n+1} T_{j,K_n+2} = n - nq_n,\\
&\sum_{i=1}^{K_n+2} N_{i1} + \sum_{i=1}^{K_n+1} S_{i1} = np_n - 1,\\
&\sum_{i=1}^{K_n+2} N_{ij} + \sum_{i=1}^{K_n+1} S_{ij} = L_n - 1 \text{ for } j = 2, \ldots, K_n + 1,\\
&\sum_{i=1}^{K_n+2} N_{i,K_n+2} + \sum_{i=1}^{K_n+1} S_{i,K_n+2} = n - nq_n.
\end{aligned}
$$

Let $\bar{G}(i-1, n)$ denote $G_n^{-1}\left(\frac{np_n + (i-1)L_n}{n}\right)$ for $i = 1, \ldots, K_n + 1$, define $\bar{G}(-1, n)$ as $X_0^*$, and $\bar{G}(K_n + 1, n)$ as $X_{K_n+2}^*$. Let $\bar{H}(j-1, n)$ denote $H_n^{-1}\left(\frac{np_n + (j-1)L_n}{n}\right)$ for $j = 1, \ldots, K_n + 1$, and define $\bar{H}(-1, n)$ as $Y_0^*$, $\bar{H}(K_n + 1, n)$ as $Y_{K_n+2}^*$. Let $\bar{g}(i-1, n)$ denote $g_n(\bar{G}(i-1, n))$ for $i = 1, \ldots, K_n + 1$, and $\bar{h}(j-1, n)$ denote $h_n(\bar{H}(j-1, n))$ for $j = 1, \ldots, K_n + 1$. Note that these definitions imply that $F_n^{(1)}(\bar{G}(i-1, n), \bar{H}(K_n + 1, n)) = \bar{g}(i-1, n)$, and $F_n^{(2)}(\bar{G}(K_n + 1, n), \bar{H}(j-1, n)) = \bar{h}(j-1, n)$.

Next, define $X_i$ as $\sqrt{n}\ \bar{g}(i-1, n)\,(X_i^* - \bar{G}(i-1, n))$ for $i = 1, \ldots, K_n + 1$, and define $X_0$ as 0 and $X_{K_n+2}$ as 0. Define $Y_j$ as $\sqrt{n}\ \bar{h}(j-1, n)\,(Y_j^* - \bar{H}(j-1, n))$ for $j = 1, \ldots, K_n + 1$, and define $Y_0$ as 0 and $Y_{K_n+2}$ as 0.

For $i = 1, \ldots, K_n + 2$ and $j = 1, \ldots, K_n + 2$, define $Q_{ij}$ as

$$
\begin{aligned}
&F_n^{(1)}(\bar{G}(i-1, n),\ \bar{H}(j-1, n))\, X_i/\bar{g}(i-1, n) \\
&+ F_n^{(2)}(\bar{G}(i-1, n),\ \bar{H}(j-1, n))\, Y_j/\bar{h}(j-1, n) \\
&- F_n^{(1)}(\bar{G}(i-2, n),\ \bar{H}(j-1, n))\, X_{i-1}/\bar{g}(i-2, n) \\
&- F_n^{(2)}(\bar{G}(i-2, n),\ \bar{H}(j-1, n))\, Y_j/\bar{h}(j-1, n) \\
&- F_n^{(1)}(\bar{G}(i-1, n),\ \bar{H}(j-2, n))\, X_i/\bar{g}(i-1, n) \\
&- F_n^{(2)}(\bar{G}(i-1, n),\ \bar{H}(j-2, n))\, Y_{j-1}/\bar{h}(j-2, n) \\
&+ F_n^{(1)}(\bar{G}(i-2, n),\ \bar{H}(j-2, n))\, X_{i-1}/\bar{g}(i-2, n) \\
&+ F_n^{(2)}(\bar{G}(i-2, n),\ \bar{H}(j-2, n))\, Y_{j-1}/\bar{h}(j-2, n)
\end{aligned}
$$

where any term containing $X_0$, $X_{K_n+2}$, $Y_0$, or $Y_{K_n+2}$ is defined to be zero. It follows from the definitions that $\sum_{j=1}^{K_n+2} Q_{ij} = X_i - X_{i-1}$, so $\sum_{i=1}^{\alpha}\sum_{j=1}^{K_n+2} Q_{ij} = X_\alpha$ for $\alpha = 1, \ldots, K_n + 2$. Similarly, $\sum_{j=1}^{\beta}\sum_{i=1}^{K_n+2} Q_{ij} = Y_\beta$ for $\beta = 1, \ldots, K_n + 2$.

Define $p_n(i, j)$ as the probability assigned by $F_n(x, y)$ to the rectangle $(\bar{G}(i-2, n) \leqslant x \leqslant \bar{G}(i-1, n),\ \bar{H}(j-2, n) \leqslant y \leqslant \bar{H}(j-1, n))$ for $i = 1, \ldots, K_n + 2$, $j = 1, \ldots, K_n + 2$. Define $W_{ij}$ as $\dfrac{N_{ij} - np_n(i, j)}{\sqrt{np_n(i, j)}}$ and $Z_{ij}$ as $W_{ij} - \dfrac{Q_{ij}}{\sqrt{p_n(i, j)}}$, for $i = 1, \ldots, K_n + 2$ and $j = 1, \ldots, K_n + 2$. Then we have

$$
\sum_{j=1}^{K_n+2}\sum_{i=1}^{K_n+2} \sqrt{p_n(i, j)}\, Z_{ij} = \frac{-2(K_n + 1)}{\sqrt{n}},
$$

$$
\sum_{i=1}^{\alpha}\sum_{j=1}^{K_n+2} \sqrt{p_n(i,j)}\, Z_{ij} = \frac{-\sum_{i=1}^{\alpha}\sum_{j=1}^{K_n+1} T_{ji} - \alpha}{\sqrt{n}} - X_\alpha \qquad \text{for } \alpha = 1, \ldots, K_n + 1, \tag{1.2}
$$

$$
\sum_{j=1}^{\beta}\sum_{i=1}^{K_n+2} \sqrt{p_n(i,j)}\, Z_{ij} = \frac{-\sum_{j=1}^{\beta}\sum_{i=1}^{K_n+1} S_{ij} - \beta}{\sqrt{n}} - Y_\beta \qquad \text{for } \beta = 1, \ldots, K_n + 1.
$$

Let $g_n^*$ denote $\min\{g_n(x)\colon G_n^{-1}(p_n) \leqslant x \leqslant G_n^{-1}(q_n)\}$, $h_n^*$ denote $\min\{h_n(y)\colon H_n^{-1}(p_n) \leqslant y \leqslant H_n^{-1}(q_n)\}$, $b_n$ denote $\min\{g_n^*, h_n^*\}$, and $c_n$ denote $\min\{f_n(x, y)\colon G_n^{-1}(p_n) \leqslant x \leqslant G_n^{-1}(q_n), H_n^{-1}(p_n) \leqslant y \leqslant H_n^{-1}(q_n)\}$.

We assume the following:

(1.3) $f_n(x, y)$, $F_n^{(1)}(x, y)$, $F_n^{(2)}(x,y)$, $\left|\frac{\partial}{\partial x} g_n(x)\right|$, $\left|\frac{\partial}{\partial y} h_n(y)\right|$, $\left|\frac{\partial}{\partial x} f_n(x,y)\right|$, and $\left|\frac{\partial}{\partial y} f_n(x,y)\right|$ are all bounded, for all values of $x,y$, and $n$.

(1.4) $\min_{(i,j)} p_n(i,j) = 0\left(\frac{1}{K_n^2}\right)$, $\max_{(i,j)} p_n(i,j) = 0\left(\frac{1}{K_n^2}\right)$.

(1.5) There exists a finite value $\bar{b}$ so that $b_n < \bar{b}$ for all sufficiently large $n$.

(1.6) $\frac{K_n^{15}}{b_n^3\sqrt{n}}$ approaches zero as $n$ increases.

(1.7) $\frac{K_n^3}{c_n b_n\sqrt{n}}$ approaches zero as $n$ increases.

Denote by $\bar{f}_n$ the joint pdf for the random variables $\{\bar{S}_{ij}, \bar{T}_{ji}, \bar{X}_i, \bar{Y}_j, \bar{W}_{ij}\}$.

Now we construct an "artificial" joint pdf for the random variables $\{\bar{S}_{ij}, \bar{T}_{ji}, \bar{X}_i, \bar{Y}_j, \bar{W}_{ij}\}$, as follows. The $2(K_n + 1) + 1$ sets of random variables $\{S_{i1}, \ldots, S_{i,K_n+2}\}$ for $i = 1, \ldots, K_n + 1$; $\{T_{j1}, \ldots, T_{j,K_n+2}\}$ for $j = 1, \ldots, K_n + 1$; and $\{Z_{ij}: i = 1, \ldots, K_n + 2; \quad j = 1, \ldots, K_n + 2; (i,j) \neq (K_n + 2, K_n + 2)\}$ are independent of each other. The joint distribution of $\{S_{i1}, \ldots, S_{i,K_n+2}\}$ is given as follows: each $S_{ij}$ is 0 or 1, $\sum_{j=1}^{K_n+2} S_{ij} = 1$, and

$$P(S_{ij} = 1) = \frac{[F_n^{(1)}(\bar{G}(i-1,n), \bar{H}(j-1,n)) - F_n(1)(\bar{G}(i-1,n), \bar{H}(j-2,n))]}{\bar{g}(i-1, n)}$$

The joint distribution of $\{T_{j1}, \ldots, T_{j,K_n+2}\}$ is given as follows: each $T_{ji}$ is 0 or 1, $\sum_{i=1}^{K_n+2} T_{ji} = 1$, and

$$P(T_{ji} = 1) = \frac{[F_n^{(2)}(\bar{G}(i-1,n), \bar{H}(j-1,n)) - F_n^{(2)}(\bar{G}(i-2,n), \bar{H}(j-1,n))]}{\bar{h}(j-1, n)}.$$

The $(K_n + 2)^2 - 1$ random variables $\{Z_{ij}: \quad i = 1, \ldots, K_n + 2; \quad j = 1, \ldots, K_n + 2; (i,j) \neq (K_n + 2, K_n + 2)\}$ have the following joint normal pdf:

$$(\sqrt{2\pi})^{1-(K_n+2)^2} (p_n(K_n + 2, K_n + 2))^{-\frac{1}{2}} \exp\left\{-\frac{1}{2} \sum_{j=1}^{K_n+2} \sum_{i=1}^{K_n+2} z_{ij}^2\right\}$$

where $z_{K_n+2,K_n+2}$ is given by the identity $\sum_{j=1}^{K_n+2} \sum_{i=1}^{K_n+2} \sqrt{p_n(i,j)}\, z_{ij} = 0$. For future use, we note that in this distribution, the variance of $Z_{ij}$ is $1 - p_n(i,j)$, and the covariance between $Z_{ij}$ and $Z_{i'j'}$ (where $(i,j) \neq (i',j')$) is $-\sqrt{p_n(i,j)p_n(i',j')}$. Now define $\{X_1, \ldots, X_{K_n+1}, Y_1, \ldots, Y_{K_n+1}\}$ in terms of $\{S_{ij}, T_{ji}, Z_{ij}\}$ by means of equations (1.2). Defining $Q_{ij}$ as above, define $\bar{W}_{ij}$ as $Z_{ij} + \frac{Q_{ij}}{\sqrt{p_n(i,j)}}$ for

$i \leqq K_n + 1$ and $j \leqq K_n + 1$. Then define $W_{ij}$ as the closest value to $\overline{W}_{ij}$ that makes $np_n(i,j) + \sqrt{np_n(i,j)}\, W_{ij}$ an integer (positive, negative, or zero). Finally, define $W_{i,K_n+2}$ by the equation $\sum_{j=1}^{K_n+2} \sqrt{np_n(i,j)}\, W_{ij} = \sum_{j=1}^{K_n+2} N_{ij} - n \sum_{j=1}^{K_n+2} p_n(i,j)$, $W_{K_n+2,j}$ by the equation $\sum_{i=1}^{K_n+2} \sqrt{np_n(i,j)}\, W_{ij} = \sum_{i=1}^{K_n+2} N_{ij} - n \sum_{i=1}^{K_n+2} p_n(i,j)$, and $W_{K_n+2,K_n+2}$ by the equation $\sum_{j=1}^{K_n+2} \sum_{i=1}^{K_n+2} \sqrt{np_n(i,j)}\, W_{ij} = -2(K_n + 1)$, where $\sum_{j=1}^{K_n+2} N_{ij}$ and $\sum_{i=1}^{K_n+2} N_{ij}$ are defined by the equations (1.1).

Denote by $f_n^*$ the joint pdf for the random variables $\{\overline{S}_{ij}, \overline{T}_{ji}, \overline{X}_i, \overline{Y}_j, \overline{W}_{ij}\}$ induced by the process just described.

For any measurable set $C_n$ in the space of $\{\overline{S}_{ij}, \overline{T}_{ji}, \overline{X}_i, \overline{Y}_j, \overline{W}_{ij}\}$, let $P_{\overline{f}_n}(C_n)$, $P_{f_n^*}(C_n)$ denote the probabilities assigned to $C_n$ by $\overline{f}_n$ and $f_n^*$, respectively. In the next section, we prove the following:

THEOREM: $\lim_{n\to\infty} \sup_{C_n} |P_{\overline{f}_n}(C_n) - P_{f_n^*}(C_n)| = 0$.

## 2. PROOF OF THE THEOREM.

The theorem combines the results of [1], [2], and [3], and the proof is a combination of the proofs in those papers. First we list some notation and elementary results that will be used in the proof.

The symbol "log" always means the natural logarithm. For any event $E$, $\overline{E}$ denotes its negation. $\Phi(x)$ denotes the standard normal cdf. If we state that a sequence of random variables $\{V_n\}$ is $0_p(r(n))$, it means that $\frac{V_n}{r(n)\beta(n)}$ converges stochastically to zero as $n$ increases, for an arbitrary sequence $\beta(n)$ with $\beta(n)$ approaching $\infty$ as $n$ increases. Thus, if $\{V_n\}$ is $0_p(r(n))$ and $r(n)$ approaches zero as $n$ increases, then $V_n$ converges stochastically to zero as $n$ increases.

The following equalities and inequalities are well-known:

(2.1) For any events $E_1, \ldots, E_m$, $P(E_1 \cap \ldots \cap E_m) \geqq 1 - \sum_{i=1}^{m} P(\overline{E}_i)$ (Bonferroni' s inequality).

(2.2) For any $x > 0$, $1 - \Phi(x) < \frac{1}{x\sqrt{2\pi}} e^{-\frac{x^2}{2}}$.

(2.3) For any positive integer $m$, $\log m! = \frac{1}{2} \log 2\pi + \left(m + \frac{1}{2}\right) \log m - m + \frac{\omega(m)}{m}$, where $|\omega(m)| < 1$. (Stirling's formula).

(2.4) If $V$ is any random variable with $P(V < 0) = 0$, and if $b$ is any nonrandom positive value, $P(V \leqq b) \geqq 1 - \frac{E(V)}{b}$ (Chebyshev's inequality).

(2.5) If $|x| < 1$, $\log(1 + x) = x - \frac{x^2}{2} + \frac{x^3}{3(1 + \theta x)^3}$ where $|\theta| < 1$.

LEMMA 2.1: If $V_1, \ldots, V_m$ are jointly normally distributed, with zero means, variances $\sigma_1^2, \ldots, \sigma_m^2$, respectively, and any covariances, and if $\bar{\sigma}$ denotes max $(\sigma_1, \ldots, \sigma_m)$, then for any $c > 0$, $P[\max_{i=1,\ldots,m} |X_i| \leqq c] > 1 - \frac{2m\bar{\sigma}}{c\sqrt{2\pi}} \exp\left[-\frac{c^2}{2\bar{\sigma}^2}\right]$.

PROOF: A combination of (2.1) and (2.2).

LEMMA 2.2: Under both $\bar{f}_n$ and $f_n^*$,

$$\{\max_{i=1,\ldots,K_n+1} |X_i|/\sqrt{n}\,\bar{g}(i-1,n)\} = 0_p\left(\frac{K_n}{\sqrt{n}\,b_n}\right), \text{ and } \{\max_{j=1,\ldots,K_n+1} |Y_j|/\sqrt{n}\,\bar{h}(j-1,n)\} = 0_p\left(\frac{K_n}{\sqrt{n}\,b_n}\right).$$

PROOF: We only have to give the proof for $X_i$, since the proof for $Y_j$ is the same. Under $\bar{f}_n$, we can assume the joint distribution of $\{X_i\}$ is the normal distribution described in [2], and then set $c = \frac{K_n}{\sqrt{n}\,b_n}$ in Lemma 2.1, completing the proof. Under $f_n^*$, using equations (1.2) we note that $\left|X_\alpha - \sum_{i=1}^{\alpha}\sum_{j=1}^{K_n+2}\sqrt{p_n(i,j)}\,Z_{ij}\right| \leqq \frac{2(K_n+2)}{\sqrt{n}}$, and that the variance of $\sum_{i=1}^{\alpha}\sum_{j=1}^{K_n+2}\sqrt{p_n(i,j)}\,Z_{ij}$ is less than one. The result then follows directly from Lemma 2.1.

LEMMA 2.3: Under $\bar{f}_n$, $\{\max_{i,j \leqq K_n+2} |W_{ij}|/\sqrt{np_n(i,j)}\} = 0_p\left(\frac{K_n^5}{b_n\sqrt{n}}\right)$.

PROOF: Let $\{\beta_n\}$ denote a sequence of nonrandom positive quantities such that $\lim_{n\to\infty}\beta_n = \infty$, and $\lim_{n\to\infty}\frac{\beta_n K_n}{b_n\sqrt{n}} = 0$. Let $N_{ij}^{(1)}$ denote the number of $(X_i', Y_i')$ falling in the open rectangle $\{\bar{G}(i-2,n) < x < \bar{G}(i-1,n),\ \bar{H}(j-2,n) < y < \bar{H}(j-1,n)\}$; let $N_{ij}^{(0)}$ denote the number of $(X_i', Y_i')$ falling in the open rectangle

$$\left\{\begin{array}{l} G_n^{-1}\left(\frac{np_n+(i-2)L_n}{n} + \frac{\beta_n K_n}{\sqrt{n}\,b_n}\right) < x < G_n^{-1}\left(\frac{np_n+(i-1)L_n}{n} - \frac{\beta_n K_n}{\sqrt{n}\,b_n}\right), \\ H_n^{-1}\left(\frac{np_n+(j-2)L_n}{n} + \frac{\beta_n K_n}{\sqrt{n}\,b_n}\right) < y < H_n^{-1}\left(\frac{np_n+(j-1)L_n}{n} - \frac{\beta_n K_n}{\sqrt{n}\,b_n}\right) \end{array}\right\};$$

and let $N_{ij}^{(2)}$ denote the number of $(X_i', Y_i')$ falling in the open rectangle

$$\left\{\begin{array}{l} G_n^{-1}\left(\frac{np_n+(i-2)L_n}{n} - \frac{\beta_n K_n}{\sqrt{n}\,b_n}\right) < x < G_n^{-1}\left(\frac{np_n+(i-1)L_n}{n} + \frac{\beta_n K_n}{\sqrt{n}\,b_n}\right), \\ H_n^{-1}\left(\frac{np_n+(j-2)L_n}{n} - \frac{\beta_n K_n}{\sqrt{n}\,b_n}\right) < y < H_n^{-1}\left(\frac{np_n+(j-1)L_n}{n} + \frac{\beta_n K_n}{\sqrt{n}\,b_n}\right) \end{array}\right\};$$

($G_n^{-1}(a)$ is defined as $X_0^*$ if $a \leqq 0$, and as $X_{K_n+2}^*$ if $a \geqq 1$; $H_n^{-1}(a)$ is defined as $Y_0^*$ if $a \leqq 0$, and as $Y_{K_n+2}^*$ if $a \geqq 1$).

Clearly, $N_{ij}^{(0)} \leqq N_{ij}^{(1)} \leqq N_{ij}^{(2)}$, and by Lemma 2.2 with probability approaching one as $n$ increases, we have $N_{ij}^{(0)} \leqq N_{ij} \leqq N_{ij}^{(2)}$. Now $N_{ij}^{(2)} - N_{ij}^{(0)}$ is the number of pairs $(X_i', Y_i')$ which fall in a region consisting of pieces of four strips, each strip having probability (under $\bar{f}_n$) no greater than $\frac{2\beta_n K_n}{\sqrt{n}\,b_n}$. Therefore, $E\{N_{ij}^{(2)} - N_{ij}^{(0)}\} \leqq \frac{8n\beta_n K_n}{\sqrt{n}\,b_n}$, and using (2.1) and (2.4), if $\bar{\beta}_n$ is any positive value,

$P[N_{ij}^{(2)} - N_{ij}^{(0)} \leqq \bar{\beta}_n \text{ for all } i, j] \geqq 1 - \dfrac{8n(K_n+2)^2 K_n \beta_n}{\sqrt{n} b_n \bar{\beta}_n}$. Below we will choose $\beta_n$ and $\bar{\beta}_n$ so that $\dfrac{\sqrt{n} K_n^3 \beta_n}{\bar{\beta}_n b_n}$ approaches zero as $n$ increases; then we can say that with probability (under $\bar{f}_n$) approaching one as $n$ increases, we have $|N_{ij} - N_{ij}^{(1)}| \leqq \bar{\beta}_n$ for all $i, j$.

Now define $W_{ij}^{(1)}$ as $\dfrac{N_{ij}^{(1)} - np_n(i,j)}{\sqrt{np_n(i,j)}}$. Then with probability (under $\bar{f}_n$) approaching one as $n$ increases, $\left| \dfrac{W_{ij}^{(1)}}{\sqrt{np_n(i,j)}} - \dfrac{W_{ij}}{\sqrt{np_n(i,j)}} \right| \leqq \dfrac{\bar{\beta}_n}{np_n(i,j)}$ for all $i, j$. But by [3] and Lemma 2.1, $\left\{ \max \dfrac{|W_{ij}^{(1)}|}{\sqrt{np_n(i,j)}} \right\} = 0_p\left( \dfrac{K_n \sqrt{\log K_n}}{\sqrt{n}} \right)$. Since $\beta_n$ can increase arbitrarily slowly, we can let $\bar{\beta}_n$ increase at any rate that makes $\dfrac{\sqrt{n} K_n^3}{\bar{\beta}_n b_n}$ approach zero as $n$ increases. This completes the proof.

LEMMA 2.4: Under $f_n^*$, $\{ \max_{i,j \leqq K_n+2} |W_{ij}|/\sqrt{np_n(i,j)} \} = 0_p\left( \dfrac{K_n^5}{b_n \sqrt{n}} \right)$.

PROOF: For $i$ and $j$ both less than or equal to $K_n + 1$, $W_{ij}$ is within $\dfrac{1}{2\sqrt{np_n(i,j)}}$ of $\bar{W}_{ij}$, and $\bar{W}_{ij} = Z_{ij} + \dfrac{Q_{ij}}{\sqrt{p_n(i,j)}}$. Applying Lemma 2.1 to $\{Z_{ij}\}$ and $\{Q_{ij}\}$, we immediately find that under $f_n^*$, $\left\{ \max_{i,j \leqq K_n+1} \dfrac{|W_{ij}|}{\sqrt{np_n(i,j)}} \right\} = 0_p\left( \dfrac{K_n^2 \sqrt{\log K_n}}{b_n \sqrt{n}} \right)$. Since $\sum_{i=1}^{K_n+2} \sqrt{np_n(i,j)}\, W_{ij}$ is $0_p(K_n)$, we see that

$$\left\{ \max_{i \leqq K_n+1} \frac{|W_{i,K_n+2}|}{\sqrt{np_n(i,K_n+2)}} \right\} = 0_p\left( \frac{K_n^3 \sqrt{\log K_n}}{b_n \sqrt{n}} \right) \text{ and } \left\{ \max_{j \leqq K_n+1} \frac{|W_{K_n+2,j}|}{\sqrt{np_n(K_n+2,\, j)}} \right\} = 0_p\left( \frac{K_n^3 \sqrt{\log K_n}}{b_n \sqrt{n}} \right)$$

From the definition of $W_{K_n+2,\, K_n+2}$ under $f_n^*$, it follows that $\dfrac{|W_{K_n+2,\, K_n+2}|}{\sqrt{np_n(K_n+2,\, K_n+2)}} = 0_p\left( \dfrac{K_n^5}{\sqrt{n}} + \dfrac{K_n^4 \sqrt{\log K_n}}{b_n \sqrt{n}} \right)$. This completes the proof.

LEMMA 2.5: At every point $(\vec{S}_{ij}, \vec{T}_{ji}, \vec{X}_i, \vec{Y}_j, \vec{W}_{ij})$ at which $\bar{f}_n$ is positive, we can write $\log \bar{f}_n(\vec{S}_{ij}, \vec{T}_{ji}, \vec{X}_i, \vec{Y}_j, \vec{W}_{ij})$ as $A_1(n) - \dfrac{1}{2} \sum_{j=1}^{K_n+2} \sum_{i=1}^{K_n+2} Z_{ij}^2 + \Delta_1(n)$

$$+ \sum_{i=1}^{K_n+1} \sum_{j=1}^{K_n+2} S_{ij} \log \left[ \frac{F_n^{(1)}(\bar{G}(i-1,\, n), \bar{H}(j-1,\, n)) - F_n^{(1)}(\bar{G}(i-1,\, n), \bar{H}(j-2,\, n))}{\bar{g}(i-1,\, n)} \right]$$

$$+ \sum_{j=1}^{K_n+1} \sum_{i=1}^{K_n+2} T_{ij} \log \left[ \frac{F_n^{(2)}(\bar{G}(i-1,\, n), \bar{H}(j-1,\, n)) - F_n^{(2)}(\bar{G}(i-2,\, n), \bar{H}(j-1,\, n))}{\bar{h}(j-1,\, n)} \right]$$

where $A_1(n)$ is nonrandom, and $\Delta_1(n)$ converges stochastically to zero as $n$ increases under both $\bar{f}_n$ and $f_n^*$.

PROOF: From the multinomial formula, assuming that the event $ND(n)$ occurs, the joint pdf for $\{\vec{S}_{ij}, \vec{T}_{ji}, \vec{X}_i^*, \vec{Y}_j^*, \vec{N}_{ij}\}$ is the product of the following four expressions:

$$n!$$

$$\prod_{j=1}^{K_n+2}\prod_{i=1}^{K_n+2}\frac{[F_n(X_i^*, Y_j^*) - F_n(X_{i-1}^*, Y_j^*) - F_n(X_i^*, Y_{j-1}^*) + F_n(X_{i-1}^*, Y_{j-1}^*)]^{N_{ij}}}{N_{ij}!}$$

$$\prod_{i=1}^{K_n+1}\prod_{j=1}^{K_n+2}[F_n^{(1)}(X_i^*, Y_j^*) - F_n^{(1)}(X_i^*, Y_{j-1}^*)]^{S_{ij}}$$

$$\prod_{j=1}^{K_n+1}\prod_{i=1}^{K_n+2}[F_n^{(2)}(X_i^*, Y_j^*) - F_n^{(2)}(X_{i-1}^*, Y_j^*)]^{T_{ji}}$$

if $N_{ij}$, $S_{ij}$, $T_{ji}$ are nonnegative integers satisfying the equations (1.1), and $X_0^* < X_1^* < \ldots < X_{K_n+2}^*$ and $Y_0^* < \ldots < Y_{K_n+2}^*$; the joint pdf is zero, otherwise. From this it follows that when $\bar{f}_n$ is positive, we can write $\log \bar{f}_n(\vec{S}_{ij}, \vec{T}_{ji}, \vec{X}_i, \vec{Y}_j, \vec{W}_{ij})$ as the sum of the following six expressions:

(2.6) $$\log n!$$

(2.7) $$-\sum_{j=1}^{K_n+2}\sum_{i=1}^{K_n+2}\log\,(np_n(i,j) + \sqrt{np_n(i,j)}\,W_{ij})!$$

(2.8) $$\sum_{j=1}^{K_n+2}\sum_{i=1}^{K_n+2}\left[\begin{array}{l} np_n(i,j) \\ +\sqrt{np_n(i,j)}\,W_{ij}\end{array}\right]$$

$$\times \log\left[\begin{array}{l} F_n\left(\bar{G}(i-1,n) + \dfrac{X_i}{\sqrt{n}\,\bar{g}(i-1,n)},\ \bar{H}(j-1,n) + \dfrac{Y_j}{\sqrt{n}\,\bar{h}(j-1,n)}\right) \\ -F_n\left(\bar{G}(i-2,n) + \dfrac{X_{i-1}}{\sqrt{n}\,\bar{g}(i-2,n)},\ \bar{H}(j-1,n) + \dfrac{Y_j}{\sqrt{n}\,\bar{h}(j-1,n)}\right) \\ -F_n\left(\bar{G}(i-1,n) + \dfrac{X_i}{\sqrt{n}\,\bar{g}(i-1,n)},\ \bar{H}(j-2,n) + \dfrac{Y_{j-1}}{\sqrt{n}\,\bar{h}(j-2,n)}\right) \\ +F_n\left(\bar{G}(i-2,n) + \dfrac{X_{i-1}}{\sqrt{n}\,\bar{g}(i-2,n)},\ \bar{H}(j-2,n) + \dfrac{Y_{j-1}}{\sqrt{n}\,\bar{h}(j-2,n)}\right)\end{array}\right]$$

(2.9) $$\sum_{i=1}^{K_n+1}\sum_{j=1}^{K_n+2}S_{ij}\log\left[\begin{array}{l} F_n^{(1)}\left(\bar{G}(i-1,n) + \dfrac{X_i}{\sqrt{n}\,\bar{g}(i-1,n)},\ \bar{H}(j-1,n) + \dfrac{Y_j}{\sqrt{n}\,\bar{h}(j-1,n)}\right) \\ -F_n^{(1)}\left(\bar{G}(i-1,n) + \dfrac{X_i}{\sqrt{n}\,\bar{g}(i-1,n)},\ \bar{H}(j-2,n) + \dfrac{Y_{j-1}}{\sqrt{n}\,\bar{h}(j-2,n)}\right)\end{array}\right]$$

(2.10) $$\sum_{j=1}^{K_n+1}\sum_{i=1}^{K_n+2}T_{ji}\log\left[\begin{array}{l} F_n^{(2)}\left(\bar{G}(i-1,n) + \dfrac{X_i}{\sqrt{n}\,\bar{g}(i-1,n)},\ \bar{H}(j-1,n) + \dfrac{Y_j}{\sqrt{n}\,\bar{h}(j-1,n)}\right) \\ -F_n^{(2)}\left(\bar{G}(i-2,n) + \dfrac{X_{i-1}}{\sqrt{n}\,\bar{g}(i-2,n)},\ \bar{H}(j-1,n) + \dfrac{Y_j}{\sqrt{n}\,\bar{h}(j-1,n)}\right)\end{array}\right]$$

$$(2.11) \qquad -\sum_{i=1}^{K_n+1} \log\left(\sqrt{n}\,\bar{g}(i-1,n)\right) - \sum_{j=1}^{K_n+1} \log\left(\sqrt{n}\,\bar{h}(j-1,n)\right).$$

Now we carry out the following sequence of calculations.

Apply (2.3) to each term in (2.7), and write $\log\left(np_n(i,j) + \sqrt{np_n(i,j)}\,W_{ij}\right)$ as $\log np_n(i,j) + \log\left[1 + \dfrac{W_{ij}}{\sqrt{np_n(i,j)}}\right]$, then apply equation (2.5) to $\log\left[1 + \dfrac{W_{ij}}{\sqrt{np_n(i,j)}}\right]$ for all $i, j$.

Expand $F_n\left[\bar{G}(i-1,n) + \dfrac{X_i}{\sqrt{n}\,\bar{g}(i-1,n)},\ \bar{H}(j-1,n) + \dfrac{Y_j}{\sqrt{n}\,\bar{h}(j-1,n)}\right]$ around $(\bar{G}(i-1,n), \bar{H}(j-1,n))$, treating $\dfrac{X_i}{\sqrt{n}\,\bar{g}(i-1,n)}$ and $\dfrac{Y_j}{\sqrt{n}\,\bar{h}(j-1,n)}$ as infinitesimals, and stopping with terms quadratic in these infinitesimals, for $i = 1, \ldots, K_n + 1$ and $j = 1, \ldots, K_n + 1$. This allows us to write the log in (2.8) as $\log\left[p_n(i,j) + \dfrac{Q_{ij}}{\sqrt{n}} + R_n(i,j)\right]$, where $\max_{i,j} |R_n(i,j)| = 0_p\left(\dfrac{K_n^2}{nb_n^2}\right)$. We note that $\sum_{i=1}^{K_n+2} \sum_{j=1}^{K_n+2} R_n(i,j) = 0$ and $\sum_{i=1}^{K_n+2} \sum_{j=1}^{K_n+2} Q_{ij} = 0$. Write $\log\left[p_n(i,j) + \dfrac{Q_{ij}}{\sqrt{n}} + R_n(i,j)\right]$ as $\log p_n(i,j) + \log\left[1 + \dfrac{Q_{ij}}{\sqrt{n}\,p_n(i,j)} + \dfrac{R_n(i,j)}{p_n(i,j)}\right]$ and apply equation (2.5) to this last logarithm, for all $i, j$.

Expand $F_n^{(1)}\left[\bar{G}(i-1,n) + \dfrac{X_i}{\sqrt{n}\,\bar{g}(i-1,n)},\ \bar{H}(j-1,n) + \dfrac{Y_j}{\sqrt{n}\,\bar{h}(j-1,n)}\right]$ around $(\bar{G}(i-1,n), \bar{H}(j-1,n))$, treating $\dfrac{X_i}{\sqrt{n}\,\bar{g}(i-1,n)}, \dfrac{Y_j}{\sqrt{n}\,\bar{h}(j-1,n)}$ as infinitesimals, and stopping with terms linear in these infinitesimals, for $i = 1, \ldots, K_n + 1$ and $j = 1, \ldots, K_n + 1$. Then treat the resulting logs in (2.9) as the logs in (2.8) were treated.

For (2.10), expand $F_n^{(2)}$ instead of $F_n^{(1)}$, and then proceed similarly. After carrying out the steps indicated, the proof of Lemma 2.5 follows directly from Lemmas 2.2, 2.3, and 2.4, and assumptions (1.3)-(1.7).

LEMMA 2.6: At every point $(\vec{S}_{ij}, \vec{T}_{ji}, \vec{X}_i, \vec{Y}_j, \vec{W}_{ij})$ at which $f_n^*$ is positive, we can write $\log f_n^*(\vec{S}_{ij}, \vec{T}_{ji}, \vec{X}_i, \vec{Y}_j, \vec{W}_{ij})$ as

$$\begin{aligned} & A_2(n) - \frac{1}{2} \sum_{j=1}^{K_n+2} \sum_{i=1}^{K_n+2} Z_{ij}^2 + \Delta_2(n) \\ & \quad + \sum_{i=1}^{K_n+1} \sum_{j=1}^{K_n+2} S_{ij} \log\left[\frac{F_n^{(1)}(\bar{G}(i-1,n), \bar{H}(j-1,n)) - F_n^{(1)}(\bar{G}(i-1,n), \bar{H}(j-2,n))}{\bar{g}(i-1,n)}\right] \\ & \quad + \sum_{j=1}^{K_n+1} \sum_{i=1}^{K_n+2} T_{ji} \log\left[\frac{F_n^{(2)}(\bar{G}(i-1,n), \bar{H}(j-1,n)) - F_n^{(2)}(\bar{G}(i-2,n), \bar{H}(j-1,n))}{\bar{h}(j-1,n)}\right] \end{aligned}$$

where $A_2(n)$ is nonrandom, and $\Delta_2(n)$ converges stochastically to zero as $n$ increases under both $\bar{f}_n$ and $f_n^*$.

PROOF: From the construction of $f_n^*$, we have that if $f_n^*(\vec{S}_{ij}, \vec{T}_{ji}, \vec{X}_i, \vec{Y}_j, \vec{W}_{ij})$ is positive, it is the product of four expressions:

$$\prod_{i=1}^{K_n+1} \prod_{j=1}^{K_n+2} \left[ \frac{F_n^{(1)}(\bar{G}(i-1,n), \bar{H}(j-1,n)) - F_n^{(1)}(\bar{G}(i-1,n), \bar{H}(j-2,n))}{\bar{g}(i-1,n)} \right]^{S_{ij}} \tag{2.12}$$

$$\prod_{j=1}^{K_n+1} \prod_{i=1}^{K_n+2} \left[ \frac{F_n^{(2)}(\bar{G}(i-1,n), \bar{H}(j-1,n)) - F_n^{(2)}(\bar{G}(i-2,n), \bar{H}(j-1,n))}{\bar{h}(j-1,n)} \right]^{T_{ji}} \tag{2.13}$$

$$J(n)(\sqrt{2\pi})^{1-(K_n+2)^2} (p_n(K_n+2, K_n+2))^{-\frac{1}{2}} \tag{2.14}$$

$$\int \cdots \int_{\bar{R}_n} \exp\left[ -\frac{1}{2} \sum_{j=1}^{K_n+2} \sum_{i=1}^{K_n+2} Z_{ij}^2 \right] d\bar{W}_{11} \cdots d\bar{W}_{K_n+1, K_n+1} \tag{2.15}$$

where $J(n)$ is the absolute value of the determinant of the transformation used in going from $\{\vec{Z}_{ij}\}$ to $\{\vec{X}_i, \vec{Y}_j, \vec{W}_{ij}\}$ for given values of $\{\vec{S}_{ij}, \vec{T}_{ji}\}$, and $\bar{R}_n$ is the region given by the inequalities $\left\{ W_{ij} - \frac{1}{2\sqrt{np_n(i,j)}} \leqq \bar{W}_{ij} \leqq W_{ij} + \frac{1}{2\sqrt{np_n(i,j)}} : i,j \leqq K_n + 1 \right\}$. Since the transformation used in going from $\{\vec{Z}_{ij}\}$ to $\{\vec{X}_i, \vec{Y}_j, \vec{W}_{ij}\}$ is linear, $J(n)$ does not depend on any of the variables. The proof of Lemma 2.6 is completed by treating (2.15) as the analogous expression was treated on pages 145 and 146 of [3].

LEMMA 2.7: Let $R_n^*$ denote the region in $(\vec{S}_{ij}, \vec{T}_{ji}, \vec{X}_i, \vec{Y}_j, \vec{W}_{ij})$-space where $f_n^*$ is positive and $\bar{f}_n$ is zero. Then $\lim_{n\to\infty} P_{f_n^*}(R_n^*) = 0$.

PROOF: $R_n^*$ consists of those points where at least one of the quantities $np_n(i,j) + \sqrt{np_n(i,j)}\, W_{ij}$ is negative. Thus, $P_{f_n^*}(\overline{R_n^*}) = P_{f_n^*}(np_n(i,j) + \sqrt{np_n(i,j)}\, W_{ij} \geqq 0, \text{ all } i,j)$. Using Lemma 2.4 and assumption 1.6, it follows that $P_{f_n^*}(\overline{R_n^*})$ approaches one as $n$ increases, proving Lemma 2.7.

Now define the random variable $V_n$ by the equation

$$\log \frac{\bar{f}_n(\vec{S}_{ij}, \vec{T}_{ji}, \vec{X}_i, \vec{Y}_j, \vec{W}_{ij})}{f_n^*(\vec{S}_{ij}, \vec{T}_{ji}, \vec{X}_i, \vec{Y}_j, \vec{W}_{ij})} = A_1(n) - A_2(n) + V_n.$$

It follows from Lemmas 2.5, 2.6, and 2.7 that $V_n$ converges stochastically to zero as $n$ increases, under both $\bar{f}_n$ and $f_n^*$. This implies that there are sequences $\{\epsilon_n\}, \{\delta_n\}$ of positive nonrandom quantities, with $\lim_{n\to\infty} \epsilon_n = 0$ and $\lim_{n\to\infty} \delta_n = 0$, such that the region $R(n)$ where $|V_n| \leqq \epsilon_n$ has probability at least $1 - \delta_n$ under both $\bar{f}_n$ and $f_n^*$. Then we have

$$1 - \delta_n \leqq \int_{R(n)} \bar{f}_n \leqq 1 \tag{2.16}$$

$$\int_{R(n)} \bar{f}_n = e^{A_1(n) - A_2(n)} \int_{R(n)} f_n^* e^{V_n} \tag{2.17}$$

$$(1 - \delta_n) e^{-\epsilon_n} \leqq \int_{R(n)} f_n^* e^{V_n} \leqq e^{\epsilon_n} \tag{2.18}$$

and from these three inequalities,

$$(1 - \delta_n) \leqq e^{A_1(n) - A_2(n)} e^{\epsilon_n} \tag{2.19}$$

$$e^{A_1(n) - A_2(n)} (1 - \delta_n) e^{-\epsilon_n} \leqq 1. \tag{2.20}$$

It follows from these two last inequalities that $A_1(n) - A_2(n)$ converges to zero as $n$ increases. By the reasoning used in [2] or [3], this completes the proof of the Theorem.

## 3. FURTHER DISCUSSION OF THE THEOREM

The theorem can be generalized in various ways. The subscripts of the chosen order statistics $X'_{[np_n+(i-1)L_n]}$ and $Y'_{[np_n+(j-1)L_n]}$ can be replaced by subscripts which are not evenly spaced, and different spacings can be used for the $X$'s and the $Y$'s. Also, assumption (1.4) is convenient, but not essential: it was not made in [3].

The Theorem was proved for the bivariate case for convenience. It is easy to see how it can be proved for cases of higher dimensions. For example, suppose we observe $n$ IID triplets $(X'_1, Y'_1, Z'_1), \ldots, (X'_n, Y'_n, Z'_n)$. We define $X^*_i, Y^*_i, Z^*_i$ for $i = 1, \ldots, K_n + 1$, as above. $N_{ijk}$ is defined as the number of triplets $(X'_i, Y'_i, Z'_i)$ in the open cube in $(x, y, z)$-space defined by $(X^*_{i-1} < x < X^*_i,\ Y^*_{j-1} < y < Y^*_j,\ Z^*_{k-1} < z < Z^*_k)$. The indicator variables $\{S_{ij}\}$, $\{T_{ji}\}$ are now replaced by the following indicator variables: $S_{ijk} = 1$ if the triplet $(X', Y', Z')$ from which $X^*_i$ came is such that $Y^*_{j-1} < Y' < Y^*_j$ and $Z^*_{k-1} < Z' < Z^*_k$ and $S_{ijk} = 0$, otherwise; $T_{ijk} = 1$ if the triplet $(X', Y', Z')$ from which $Y^*_j$ came is such that $X^*_{i-1} < X' < X^*_i$ and $Z^*_{k-1} < Z' < Z^*_k$ and $T_{ijk} = 0$, otherwise; $U_{ijk} = 1$ if the triplet $(X', Y', Z')$ from which $Z^*_k$ came is such that $X^*_{i-1} < X' < X^*_i$ and $Y^*_{j-1} < Y' < Y^*_j$, and $U_{ijk} = 0$, otherwise. The rest of the development follows the two-dimensional case almost exactly. The necessary modifications of the conditions (1.3)-(1.7) are easily made by following the computations in Lemma 2.5.

Finally, we discuss the event $ND(n)$. The assumption that $ND(n)$ occurs is used only to guarantee that the equations (1.1) hold. Since these equations hold by construction under $f^*_n$, we only investigate $ND(n)$ under $\bar{f}_n$.

In the special case where $X'_i$ and $Y'_i$ are independent, $P_{\bar{f}_n}[ND(n)]$ is equal to $\binom{n-K_n-1}{K_n+1} \Big/ \binom{n}{K_n+1}$, and this approaches one as $n$ increases if $\dfrac{K_n^2}{n}$ approaches zero as $n$ increases.

In cases where $X'_i$ and $Y'_i$ are not independent, we proceed as follows. Let $M_n(\alpha, \beta)$ denote the total number of pairs $(X'_i, Y'_i)$, with $X'_i < X^*_\alpha$ and $Y'_i < Y^*_\beta$. Define $S_1(\alpha, \beta)$ as 1 if the $Y'$ originally associated with $X^*_\alpha$ falls below $Y^*_\beta$, and zero otherwise; define $S_2(\alpha, \beta)$ as 1 if the $Y'$ originally associated with $X^*_\alpha$ is equal to $Y^*_\beta$, and as 0 otherwise; define $S_3(\alpha, \beta)$ as 1 if the $Y'$ associated with $X^*_\alpha$ is greater than $Y^*_\beta$, and zero otherwise. Define $T_1(\beta, \alpha)$ as 1 if the $X'$ originally associated with $Y^*_\beta$ falls below $X^*_\alpha$, and zero otherwise; define $T_2(\beta, \alpha)$ as 1 if the $X'$ originally associated with $Y^*_\beta$ is equal to $X^*_\alpha$, and zero otherwise; define $T_3(\beta, \alpha)$ as 1 if the $X'$ originally associated with $Y^*_\beta$ is greater than $X^*_\alpha$, and zero otherwise. Thus, $S_1(\alpha, \beta) + S_2(\alpha, \beta) + S_3(\alpha, \beta) = 1$, $T_1(\beta, \alpha) + T_2(\beta, \alpha) + T_3(\beta, \alpha) = 1$, and $S_2(\alpha, \beta) = T_2(\beta, \alpha)$ with probability one. It is easy to write the joint pdf for $M_n(\alpha, \beta)$, $X^*_\alpha$ $Y^*_\beta$, $S_1(\alpha, \beta)$, $S_2(\alpha, \beta)$, $S_3(\alpha, \beta)$, $T_1(\beta, \alpha)$, $T_2(\beta, \alpha)$, $T_3(\beta, \alpha)$, by using the multinomial formula. From this joint pdf, and the fact that under our assumptions

$$\left| \frac{M_n(\alpha,\beta)}{n} - F_n\left(\bar{G}(\alpha-1,n), \bar{H}(\beta-1,n)\right)\right|, \quad |X^*_\alpha - \bar{G}(\alpha-1,n)|, \text{ and } |Y^*_\beta - \bar{H}(\beta-1,n)|$$

all converge stochastically to zero as $n$ increases, uniformly for $\alpha, \beta = 1, \ldots, K_n + 1$, it is easy to show that $P_{\bar{f}_n}(S_2(\alpha, \beta) = 1) - \dfrac{f_n(\bar{G}(\alpha-1,n), \bar{H}(\beta-1,n))}{n\,\bar{g}(\alpha-1,n)\,\bar{h}(\beta-1,n)}$ is $o\left(\dfrac{1}{n}\right)$ uniformly for $\alpha, \beta = 1, \ldots, K_n + 1$. Thus, $\sum_{\beta=1}^{K_n+1} \sum_{\alpha=1}^{K_n+1} P_{\bar{f}_n}(S_2(\alpha, \beta) = 1) = 0\left(\dfrac{K_n^2}{nb_n^2}\right)$ and approaches zero as $n$ increases, by assumption (1.6). Since the event $\{S_2(\alpha, \beta) = 1\}$ is the same as the event that $X^*_\alpha$ and $Y^*_\beta$ were originally observed as a pair, if follows from (2.1) that $P_{\bar{f}_n}(ND(n))$ approaches one as $n$ increases, under our assumptions.

## 4. A TEST OF INDEPENDENCE

Suppose that the problem is to test the hypothesis that $X_i'$ and $Y_i'$ are independent. Since a strictly monotonic transformation applied to each $X_i'$ does not affect independence of $X_i'$ and $Y_i'$ a reasonable test of independence should not depend on the values of the order statistics. We will construct a test of independence based on $\{\vec{N}_{ij}\}$.

Choose $p_n = 1 - q_n = \frac{L_n}{n}$. Then, if the hypothesis of independence is true, $p_n(i, j) = \frac{1}{(K_n+2)^2}$ for all $i$ and $j$. Also, it is easily shown that under the hypothesis of independence, Variance $\{\overline{W}_{ij}\} = \left(\frac{K_n+1}{K_n+2}\right)^2$, Covariance $\{\overline{W}_{ij}, \overline{W}_{ij}\}$ = Covariance $\{\overline{W}_{ij}, \overline{W}_{i'j}\} = -\frac{K_n+1}{(K_n+2)^2}$, Covariance $\{\overline{W}_{ij}, \overline{W}_{i'j'}\} = \frac{1}{(K_n+2)^2}$ if $i \neq i'$ and $j \neq j'$. These formulas can also be assumed to hold if $(K_n+2)^2 p_n(i, j)$ approaches one as $n$ increases, uniformly in $i$ and $j$, and all asymptotic probability calculations will be correct.

Define $Q(n)$ as $\sum_{j=1}^{K_n+2} \sum_{i=1}^{K_n+2} \left[\frac{(K_n+2)N_{ij} - n/(K_n+2)}{\sqrt{n}}\right]^2$, and note that $Q(n)$ is observable. Then, if $(K_n+2)^2 p_n(i, j)$ approaches one as $n$ increases, uniformly in $i$ and $j$, the asymptotic distribution of $Q(n)$ is normal, with mean $(K_n+1)^2 + n \sum_{j=1}^{K_n+2} \sum_{i=1}^{K_n+2} \left[(K_n+2)p_n(i, j) - \frac{1}{K_n+2}\right]^2$ and variance $2(K_n+1)^2$. Then we have the following test of independence of $X_i'$ and $Y_i'$: Reject the hypothesis if $\frac{Q(n) - (K_n+1)^2}{\sqrt{2(K_n+1)^2}} > \Phi^{-1}(1-\alpha)$, where $\alpha$ is the desired asymptotic level of significance. The asymptotic power of this test is

$$1 - \Phi\left[\Phi^{-1}(1-\alpha) - \frac{n}{\sqrt{2(K_n+1)^2}} \sum_{j=1}^{K_n+2} \sum_{i=1}^{K_n+2} \left\{(K_n+2)p_n(i, j) - \frac{1}{K_n+2}\right\}^2\right].$$

## REFERENCES

[1] Weiss, L., "On the Asymptotic Joint Normality of Quantiles from a Multivariate Distribution," Journal of Research of the National Bureau of Standards, *68B*, 65-66 (1964).

[2] Weiss, L., "Statistical Procedures Based on a Gradually Increasing Number of Order Statistics," Communications in Statistics, *2*, 95-114 (1973).

[3] Weiss, L., "The Normal Approximation to the Multinomial with an Increasing Number of Classes," Naval Research Logistics Quarterly, *23*, 139-149 (1976).

# MINIMIZING COSTS OF SEQUENTIAL AND PARALLEL PERSONNEL TESTING PROGRAMS

Paul D. Feigin and Gedaliahu H. Harel

Faculty of Industrial Engineering and Management
Technion-Israel Institute of Technology
Haifa, Israel

ABSTRACT

We consider groups of tests for personnel selection purposes in which each test has a known a priori probability of being failed, such failure resulting in outright rejection and termination of testing. Each test has a fixed cost and given duration. We consider the minimization of the total expected cost due to both the fixed costs and the delay costs when the tests may be conducted sequentially or in parallel. In the latter situation, a heuristic algorithm is proposed and illustrated.

## 1. INTRODUCTION AND MOTIVATION

Consider the problem of determining the suitability of a job candidate for a particular position. The personnel manager bases his judgment on an $n$-dimensional profile—each dimension being a score on a particular aspect of relevance to the candidate's acceptability. Our interest is in determining the optimal order of assembling this profile given that any single very low score can cause the candidate's rejection no matter what the nature of the remainder of his profile. Moreover, the various scores are based on tests which are assumed to be independent so that our problem differs from that of a correlated battery of tests, all of which are used to predict the individual's success (see, for example, Cronbach and Glazer [3]). Our assumptions are reasonable when the dimensions represent aspects as diverse as health, security rating, intelligence, experience, vocational interest, and personality.

For each test there is a predetermined critical score and a candidate whose score is below this level is rejected outright and no further testing is required. We assume that the proportion of the population who may fail in this way is known for each test. Conducting each test involves a fixed cost as well as a waiting time for the receiving of the results. The object is to arrange the tests so as to minimize the expected costs of testing and delay in reaching a decision. The decision will either be:

(i) rejection due to a "failure," or

(ii) that based on the complete profile.

We are not concerned here with how the personnel manager is to relate his decision to the candidate's profile in cases when the latter passes each test—he may even automatically accept all such candidates.

It is our belief that in many employment decision problems, the situation described above is more realistic than the common formulation of sequential testing. In the latter case the suggestion has been to place "the selection device that has the highest correlation with job success (validity)" first (Beach [2]

p. 239). This approach completely ignores the role costs of testing must play in determining the optimal sequence. Moreover, rather extensive research has to be done on the correlations between the various tests and the success criterion for each different job. Our approach provides a set ordering based only on two types of information—the costs associated with each test and the probability of failing each test. This sort of information may often be available from general sources and does not require new experimental information.

We also consider the situation were one does not necessarily wait for the results from a completed test before conducting a new one. This possibility leads us into the realm of parallel testing procedures and it is here that the major contributions are made based on heuristic algorithms.

## 2. SEQUENTIAL TESTING

The personnel manager wishes to obtain the candidate's score on $n$ tests $T_1, \ldots, T_n$ in order to consider his job suitability. Each test has a veto failure level: a score below this level (a failure) automatically precludes the applicant from taking the job. In the interest of saving time one would suggest starting all tests immediately; on the other hand, if there is a good chance of a candidate failing a particular test it may be worthwhile waiting for the results on this test since there is a good chance of saving the fixed costs associated with further tests.

For the case in which testing can only be performed sequentially we formalize the discussion by considering the following assumptions:

ASSUMPTION 1: The nature of the tests is such that the event of passing a particular test is independent of that for any other test.

ASSUMPTION 2: The a priori probability $p_i$ that the candidate may fail test $T_i$ is known for each $i = 1, 2, \ldots, n$.

ASSUMPTION 3: The cost of waiting to make a decision based on the test results is $\gamma$ units per unit time. (Decision is made according to (i) or (ii) of Section 1.)

ASSUMPTION 4: Tests may not be run simultaneously, nor may a test be started while the results on another are pending.

Assumption 2 involves knowledge of the probability that the candidate will fail each of the tests. These numbers can either be determined from general population proportions or from proportions appropriate to the subclass to which the applicant belongs. In other words, after a preliminary interview, or from the job application form, the personnel manager may be able to determine to which broad category the candidate belongs and thus estimate the appropriate failure probabilities. If these probabilities differ from category to category then so may the optimal sequencing of tests.

Assumption 4 will be reviewed in Section 3.

Denote by $c_i$ the fixed cost associated with conducting test $T_i$, and let $d_i$ be the duration of this test. Duration, in some contexts, may refer to the actual time span of the test or, it may refer to the waiting time for the results to be obtained. In the latter case, the actual physical involvement of the candidate in the test will be considered negligible (see Section 3) as far as time is concerned.

Under Assumption 4 we can associate the cost

$$a_i = c_i + \gamma\, d_i$$

with test $T_i$, since there is no parallel testing.

RESULT 1: Under Assumptions 1, 2, 3 and 4, the optimal testing sequence is determined by quantities $\{a_i/p_i\}$—the test with the smallest value is conducted first and so on.

PROOF: For a given ordering, say $T_1, T_2, \ldots, T_n$, the expected cost is given by

$$E(C) = a_1 + (1-p_1)a_2 + (1-p_1)(1-p_2)a_3 + \ldots + (1-p_1)\ldots(1-p_{n-1})a_n. \tag{1}$$

Suppose we change the order of tests $T_i$ and $T_{i+1}$. The relevant contribution to $E(C)$ in (1) is

$$(1-p_1)(1-p_2)\ldots(1-p_{i-1})(a_i + (1-p_i)a_{i+1})$$

whereas under the new order, it is

$$(1-p_1)(1-p_2)\ldots(1-p_{i-1})(a_{i+1} + (1-p_{i+1})a_i),$$

the other terms remaining unchanged. From here it is clear that $T_i$ should appear before $T_{i+1}$ if $a_i/p_i < a_{i+1}/p_{i+1}$, and the result follows straightforwardly.

The above proof is a simple application of the "adjacent pairwise interchange method" (see, for example, Baker [1]).

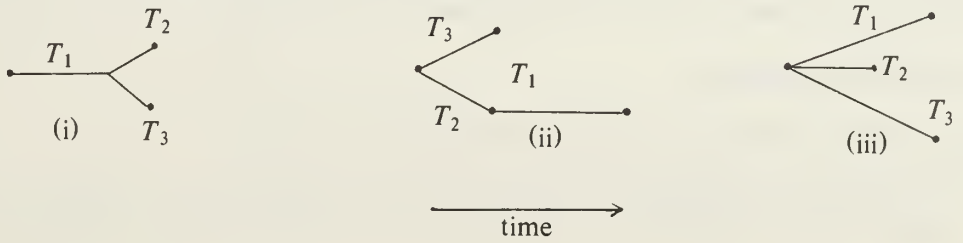
## 3. PARALLEL TESTING

The restriction that no test be commenced before the previous results are known is an artificial restriction for most personnel selection situations. If time, and hence money, can be saved by running several tests simultaneously then this option should also be investigated. If it is financially important for a firm to fill its vacancy (or vacancies) as quickly as possible (that is, $\gamma$ is large), then it will become more cost efficient to run several tests in parallel. Depending on the value of $\gamma$ and the set up costs, this does not necessarily mean that all the tests should be started at the outset. We therefore turn to the problem of approximating the optimal testing program.

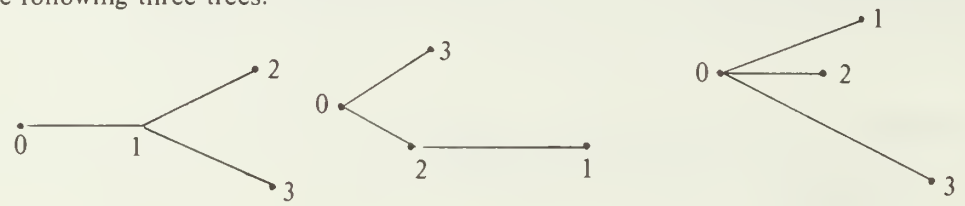
### A. The Size of the Problem

Here we drop Assumption 4 and allow several tests to be "run" at the same time—for the personnel selection situation a test is "running" as long as its result is still pending—the actual active part of the test, as far as the candidate is concerned, is considered negligible. In other words, we consider situations in which processing and making the results available to the employer take considerably longer than the administration of the tests themselves. People involved in personnel selection are well aware of these procedural delays. To get some idea of the size of the problem, we first consider how many possible arrangements of the $n$ tests exist when an arbitrary number may be conducted at the same time. The optimality problem is significantly simplified by the observation that there is no point starting a particular test when no other test is finishing—the fixed costs being incurred in any case—and, in the interest of minimizing delay, one might as well advance the start until the end of the most recently completed test.

The various possible configurations can be represented by trees. For example, for $n = 3$ tests, consider three possible configurations

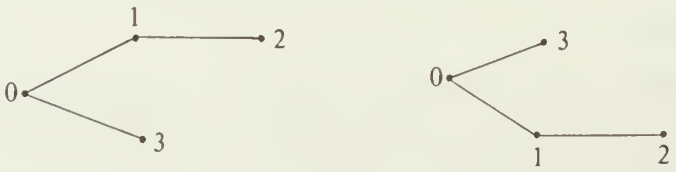


The first configuration (i) indicates that test $T_1$ is started first, followed simultaneously by $T_2$ and $T_3$. We can use the length of lines to indicate the relative durations and label nodes by the number of the test just completed, using 0 to indicate the time origin. The above configurations are then represented by the following three trees:




The number of such trees is given by $T(n) = (n+1)^{n-1}$ which corresponds to the number of configurations of $n$ tests (see Harary, [5] chapter 11) and for $n = 3$ there are 16 such possible arrangements.

For the testing situation there may be somewhat fewer possible arrangements if the "branch distance" from the origin is the same for two different nodes. For example, if $d_1 = d_3$ then




are effectively equivalent as far as deciding when to start $T_2$.

To obtain a lower bound $A(n)$ on the number of effectively different configurations, we consider the case when all the tests have the same duration. It is not hard to see that the following recurrence relation holds:

$$A(n) = \sum_{j=0}^{n-1} \binom{n}{j} A(j), \quad A(0) = 1.$$

Some values of $T(n)$ and $A(n)$ are given below:

| $n$ | 1 | 2 | 3 | 4 | 5 | 6 |
|---|---|---|---|---|---|---|
| $T(n)$ | 1 | 3 | 16 | 125 | 1296 | 16807 |
| $A(n)$ | 1 | 3 | 13 | 75 | 541 | 4683. |

In any given situation with $n$ tests of possibly different duration times, the number of different arrangements will be between the two numbers $A(n)$ and $T(n)$.

Seeing that finding the optimal arrangement by complete enumeration quickly becomes prohibitive, we now suggest some heuristic algorithms to provide good suboptimal schedules. The efficiency of the algorithms is discussed for some simple examples—a more thorough investigation, possibly using the sampling approach of Dannenbring [4], being postponed at this stage.

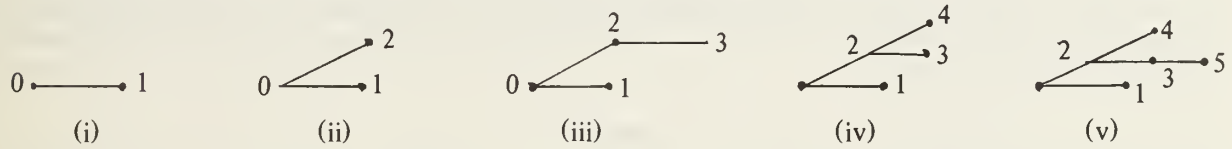
### B. Single-Step Check Algorithm

Using the notation of Section 2 we describe the algorithm as follows:

(I) Order the tests according to $(a_i/p_i)$ values as in the sequential case. Assume the resulting order is $(T_1, \ldots, T_n)$.

(II) Compare the expected costs up to the completion of all tests $T_1, \ldots, T_i$ when $T_i$ is started at the last node of the current configuration of $(T_1, T_2, \ldots, T_{i-1})$, with that incurred if $T_i$ were started simultaneously with the most recently started test. Choose the configuration with the smaller expected cost and then proceed to consider $T_{i+1}$ similarly, until all $n$ tests have been scheduled.

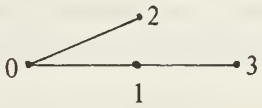
The number of comparisons required is $n - 1$; a typical succession of steps for $n = 5$ may appear as follows:




(i) (ii) (iii) (iv) (v)

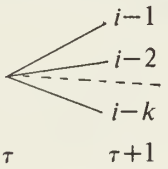
Note that moving from stage (ii) to (iii) the tree

(3)




is not considered according to (II) above. We could have considered the single step procedure which would at each stage compare trees such as (3) with that chosen (iii). However, this latter algorithm has the disadvantage of not being able to potentially produce both the purely sequential as well as the completely simultaneous schedules (unless all tests are of the same duration). It seems more important to consider single-step check procedures that can take in both extremes as against those that do not skip possible starting nodes (such as occurs when (3) is missed). If the durations $d_i$ are all equal (say 1 unit) then the above mentioned variations of (II) are equivalent.

In fact, the required comparisons can be made quite simply when $d_i = 1$, for all $i$. Note that since the tests are run in parallel and $d_i = 1$, only the fixed cost, $c_i$, of the test need be considered. Without loss of generality we take $k = 1$ here and in the sequel. At stage $i$ we have typically




where $\tau, \tau + 1$ denote units of time from the origin. Suppose the probability of reaching $\tau$ is denoted by $(1 - \pi_\tau)$. If $T_i$ is started at $\tau$ then the relevant contribution to the expected cost up to time $\tau + 1$ is

$$(1 - \pi_\tau)\,(c_{i-k} + \ldots + c_{i-1} + c_i + 1) \tag{4}$$

whereas if $T_i$ is started at time $\tau + 1$, it is

$$(1 - \pi_\tau)\,(c_{i-k} + \ldots c_{i-1} + 1) + (1 - \pi_{\tau+1})\,(c_i + 1). \tag{5}$$

Comparing (4) and (5) we decide to start $T_i$ at

$$\begin{cases} \tau \text{ if } c_i < \dfrac{(1-\pi_{\tau+1})}{1-\pi_\tau}\,(c_i + 1) \\ \tau + 1 \text{ if } c_i > \dfrac{(1-\pi_{\tau+1})}{1-\pi_\tau}\,(c_i + 1) \end{cases} \tag{6}$$

where, clearly, $\dfrac{1-\pi_{\tau+1}}{1-\pi_\tau} = (1-p_{i-k}) \ldots (1-p_{i-1})$.

For this situation we see that the comparisons can be made quickly and sequentially with a minimum of data to be stored.

The systematic application of the algorithm when the $d_i$ are different will now be described. Suppose we are at stage $i$ and wish to determine where to start test $T_i$. Suppose that $\lambda$ denotes the time of the last node so far and $\nu$ the time of the most recent start (i.e., the time at which $T_{i-1}$ is commenced). Define the following quantities:

$C_\lambda$ — expected cost up to time $\lambda$ ($T_i$ excluded);

$C_\nu$ — expected cost up to time $\nu$;

$r$ — number of tests commencing at $\nu$ (including $T_i$);

$s$ — number of tests commencing or in progress at time $\nu$;

$\nu_1, \nu_2, \ldots, \nu_s$ — indices of tests commencing or in progress at time $\nu$;

$R_1 < R_2 < \ldots < R_s$ — time left to completion of tests $\nu_1, \ldots, \nu_s$, respectively;

$K = \sum_{j=1}^{r} c_{i-j+1}$ — fixed costs of starts at time $\nu$;

$1-\pi_\nu$ — probability of getting through all tests up until time $\nu$.

Note that the index $i$ is among the indices $\nu_1, \ldots, \nu_s$.

Now if $T_i$ is started at time $\lambda$ then the new expected cost is

$$C'_\lambda = C_\lambda + \left\{\prod_{j=1}^{i-1}(1-p_j)\right\}(c_i + d_i) \tag{7}$$

with the remainder of the updating as follows:

$$\lambda' = \lambda + d_i,\ \nu' = \lambda,\ C'_\nu = C_\lambda,\ r' = s' = 2,\ \{\nu'_1, \nu'_2\} = \{i, i+1\},$$

$$R'_1 = \min(d_i, d_{i+1}),\ R'_2 = \max(d_i, d_{i+1}),\ K' = c_i + c_{i+1},$$

$$1-\pi_\nu = \prod_{j=1}^{i-1}(1-p_j).$$

If $T_i$ is started at time $\lambda$ then the new expected cost is

$$C'_\lambda = C_\nu + (1-\pi_\nu)\{K + p_{\nu_1}R_1 + (1-p_{\nu_1})p_{\nu_2}R_2 + \ldots + (1-p_{\nu_1})\ldots(1-p_{\nu_{s-1}})R_s\} \tag{8}$$

with the appropriate updating determined by:

$$\lambda' = \nu + R_s,\ \nu' = \nu,\ C'_\nu = C_\nu,\ r' = r+1,\ s' = s+1,\ K' = K + c_{i+1},\ \pi'_\nu = \pi_\nu$$

and $d_{i+1}$ being inserted into its correct place among $R_1, \ldots, R_s$ to give the new indices $\nu'_1, \ldots, \nu'_{s+1}$ and residual times $R'_1, \ldots, R'_{s+1}$.

At each stage $i$ the smaller of (7) and (8) determines the appropriate scheduling of $T_i$ and the appropriate updated quantities to proceed to the next stage.

### C. Multistep Check Algorithms

At the expense of considerably increased computational effort the single-step check algorithm can be extended by considering starting test $T_i$ at several nodes in the existing configuration at stage $i$. There are at most $i$ possible starting nodes so that if each one of these was investigated at each stage $i$, at most $n(n+1)/2$ different configurations would be considered.

For example, one might start by considering a two-step check procedure which would involve comparing the merit of starting, at stage $i$, $T_i$ at either the last node or one of the last two *starting* times. The algorithm would require storing considerably more information at each stage. However, there is no guarantee that the final schedule will be superior to that produced by the single-step check procedure.

Indeed, due to the serious interactive effect of parallel tests on the cost structure, it is not possible to conclude that any multistep procedure will always do better than a one-step procedure. This is because on adding $T_n$, we may be better off with some $T_i$ starting at a node determined by the one-step procedure than at the node to which it was allocated at stage $i$ by the multistep procedure.

We therefore do not look further at the multistep check algorithms in this paper, although their accurate description still poses an interesting and possibly important problem.

## 4. NUMERICAL EXAMPLES

EXAMPLE 1.

Data: $n = 3$

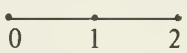
| $i$ | $c_i$ | $d_i$ | $p_i$ | $a_i$ | $a_i/p_i$ |
|---|---|---|---|---|---|
| 1 | 1 | 1 | .25 | 2 | 8 |
| 2 | 8 | 1 | .7 | 9 | 90/7 |
| 3 | 2 | 1 | .2 | 3 | 15 |

ALGORITHM: We note that $d_i = 1$; $i = 1, 2, 3$ so that we may use the criterion (6) at each stage.

(1) Starting order is 1, 2, 3.

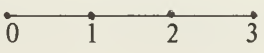
(2) Start $T_2$ at 0 if $c_2 < (1-p_1)(c_2+1) = (.75)(9) = 6.75$.

Hence, we start $T_2$ at time 1 and the configuration stands as



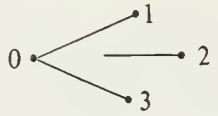

(3) Start $T_3$ at node 1 if $c_3 < (1-p_2)(c_3+1) = (.3)(3) = .9$.

Hence, we start $T_3$ at node 2 and the final configuration is




with total expected cost 9.425

The optimal configuration is actually




with a total expected cost of 9.4. We note that a two-step check procedure would have given us this configuration but that the one-step algorithm gives a very nearly optimal schedule. Here we also have an example of a case for which the optimal parallel testing schedule involves inversions of the testing order for a sequential schedule (viz. starting $T_2$ on the completion of $T_3$).

EXAMPLE 2.

Data: $n = 3$

| i | $c_i$ | $d_i$ | $p_i$ | $a_i$ | $a_i/p_i$ |
|---|---|---|---|---|---|
| 1 | 1 | 1 | .3 | 2 | 20/3 |
| 2 | 4 | 3 | .6 | 7 | 35/3 |
| 3 | 1 | 2 | .2 | 3 | 15 |

ALGORITHM: We use the algorithm outlined by equations (7) and (8). The initial order is again 1, 2, 3, so that stage (1) consists of starting $T_1$ at 0. We now tabulate the relevant quantities defined in the previous section for stages (2) and (3), explaining the calculations below.

| | $\lambda$ | $\nu$ | $r$ | $s$ | $\{\nu_j\}$ | $\{R_j\}$ | $K$ | $1-\pi_\nu$ | $C_\nu$ | $C_\lambda$ |
|---|---|---|---|---|---|---|---|---|---|---|
| (2) | 1 | 0 | 2 | 2 | 1, 2 | 1, 3 | 5 | 1 | 0 | 2 |
| (3) | 4 | 1 | 2 | 2 | 3, 2 | 2, 3 | 5 | .7 | 2 | 6.9 |

Stage 2: According to (7) $C'_\lambda = 2 + (.7)7 = 6.9$

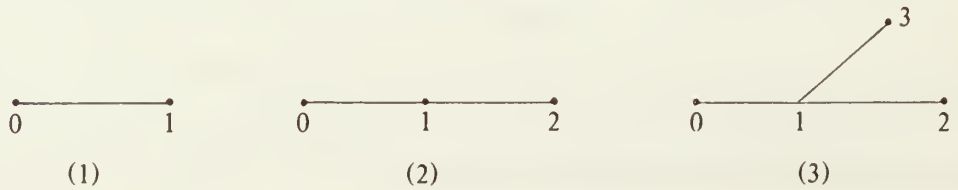
According to (8) $C'_\lambda = 0 + (1)\{5 + (.3)\,1 + .7(3)\} = 7.4$

We therefore update according to equation (7).

Stage 3: According to (7) $C'_\lambda = 6.9 + (.7)\,(.4)\,(3) = 7.74$

According to (8) $C'_\lambda = 2 + (.7)\{5 + (.2)\,2 + (.8)\,3\} = 7.46$

We do not need to update the quantities further. The tree has developed as follows:




The algorithm does produce the optimal configuration in this case, with total expected cost 7.46.

## CONCLUSION

We have proposed an approach to personnel test sequencing which is based on cost effectiveness rather than predictive ability. Once a company has decided on which tests are required to determine job

suitability, then their application may be done more efficiently by considering the sequencing algorithms discussed above.

It is clear that a major area of application lies in recruiting for specialized jobs in the government or the armed forces. Here, investigating the various facets such as security, health, intelligence and personality typically involve high cost and delays. Moreover, information is commonly available on the a priori probability that an applicant will meet the required standards in each area. We suggest that the ideas presented here will help make the screening of applicants more efficient, either by implementing the algorithms herein described or by encouraging the development of related ones.

## REFERENCES

[1] Baker, K.R., *Introduction to Sequencing and Scheduling* (John Wiley and Sons, New York, 1974).
[2] Beach, D.S., *Personnel: The Management of People at Work* (MacMillan Publishing Co., New York, 1975).
[3] Cronbach, L.J. and C.C. Gleser, *Psychological Tests and Personnel Decisions* (University of Illinois Press, Urbana, 1957).
[4] Dannenbring, D.C., "Estimating Optimal Solutions for Large Combinatorial Problems," Management Science, *23*, 1273-1283 (1977).
[5] Harary, F., *A Seminar on Graph Theory* (Holt, Rinehardt and Winston, New York, 1967).

# THE VARIANCE OF THE COVERAGE OF A RANDOMLY LOCATED AREA TARGET BY A SALVO OF WEAPONS

Gerhard Schroeter

*Industrieanlagen-Betriebsgesellschaft mbH*
*Ottobrunn, Germany*

ABSTRACT

We derive formulas for the variance of that proportion of the value of a randomly located, circularly symmetric area target that is destroyed by $N$ independently fired weapons of identical type whose damage functions are circularly symmetric about the respective impact points. The probability density functions of the target center location and of the weapon impact points are also circularly symmetric. The general results are specialized to uniform and Gaussian functions. In the latter case a closed-form solution (triple integral) for the variance of the coverage is derived. Similar to some well-known results on expected coverage, this expression for the variance of the coverage can be easily evaluated by numerical quadrature. Numerical results are given which indicate the target coverage variability caused by the combined effects of random target-locating errors and weapon impact point fluctuations.

## INTRODUCTION

Two-dimensional coverage problems arise frequently when assessing the effectiveness of artillery or air-to-ground weapons. The surveying papers of Eckler [5], Eckler and Burr [6], Guenther and Terragno [16] show that there are many papers concerned with the expected coverage of an area target by multiple independently fired weapons. This reflects the importance of knowing the average effectiveness of weapon systems. However, in military requirements another criterion is often used: The measure of effectiveness for a salvo of $N$ rounds is the proportion $\epsilon$ of the total value of an area target that is destroyed with a probability of at least $\gamma$. Expected coverage, obviously, does not supply such information, even though, of course, a higher expectation of destroyed value generally increases the confidence level $\gamma$ for any given $\epsilon$ (or vice versa). What would be required in the case of an $\epsilon$-$\gamma$-criterion is not only the expectation of the destroyed value or coverage, but the distribution of the coverage. In sharp contrast to the bulk of papers on expected coverage, only very little literature is available on the distribution problem. Domb [4], Siegel [34] and Votaw [38] have calculated distributions of the coverage for arbitrary numbers of covering objects $N \geqslant 1$; they, however, treat only one-dimensional coverage problems (line segments and circular arcs), and their methods do not seem applicable to area problems. Solomon [35] derives in two dimensions the distribution of the destroyed value of an area target by one weapon only and gives some upper and lower bounds for $\gamma$ (given $\epsilon$) in the case of two weapons. Several authors have derived expressions for the higher moments (especially the second moment) of that fraction of a fixed geometrical target figure or body which is covered by $N$ randomly thrown figures or bodies (Ailam [1], Bronowski and Neyman [2], Garwood [13], Greenberg [15], Moran and Fazekas de St. Groth [24], Robbins [25] and [26], Santaló [27]). The principal result obtained in this connection is due to Robbins [25]. Others (Cooke [3], Edens [7], Flatto [8], Flatto and Konheim [9], Flatto and Newman [10], Gilbert [14], Holst [18] and [19], Kaplan [20] and [21], Miles [23], Shepp [32], Siegel [33], Steutel [36], Stevens [37]) estimate the probability that the randomly thrown figures cover the fixed target figure completely, or estimate the number of figures required to cover the target figure completely.

From all these results, those on the second moment of the coverage in the space of two or more dimensions are the most interesting with respect to weapon analyses. Although the second moment (or the variance) does not supply the precise information that is required for the $\epsilon$-$\gamma$-criterion, it is nevertheless useful, since Tchebyshev's inequality can in some cases yield a lower bound for $\epsilon$ (given $\gamma$) or $\gamma$ (given $\epsilon$). The papers cited, however, are confined to one situation; this limitation results partly from modeling the coverage problem as an overlap-of-geometrical-figures problem. In weapon analyses terms, it is: (1) The only targets considered are those with sharp edges and uniform value density. (2) The only weapons (randomly thrown figures) are those with zero-one damage functions; complete destruction occurs within the weapon effects area, and no destruction without. (3) The weapon impact points are assumed to be distributed uniformly within a specified area. For most applications, these three assumptions could be accepted as more or less good approximations of the real world. However, one limitation is a serious drawback, which reduces the usefulness of the former results on the variance of the coverage: (4) The target position is not a random variable, or in our words: The target location error is zero with probability one.

It is the objective of this paper to derive expressions for the variance of the coverage in more general situations so that limitations (1) to (4) are eliminated. First, we generalize Robbins's [25] main result on the moments of the coverage of geometrical figures to randomly located targets with general value density function. Second, we consider a coverage problem in which all underlying functions are Gaussian, and derive a closed-form solution (triple integral) for the variance of the coverage. In this Gaussian case the standard deviation of the coverage can be determined in little computer time by numerical quadrature so that a parametric analysis can be made. Numerical results will be presented which clearly indicate the variability of target coverage caused by the combined effects of random weapon impact point and target location fluctuations. As a side result, we generalize a formula of Garwood [13] and Santaló [27] for the variance of the coverage of one circle by $N$ other circles.

The concepts of the Gaussian target and the Gaussian damage function have been considered in many papers (see, e.g., Fraser [11] and [12], Eckler [5], Eckler and Burr [6]), dealing with problems on expected coverage; the main advantage of this model—beyond its intuitive appeal as an approximation of real world situations—is that many problem solutions can be given in closed form which is not possible when corresponding problems with other functions are considered. This same property will ultimately be used to derive the triple integral expression for the variance of the coverage of the normally located Gaussian target by warheads with Gaussian damage functions and normally distributed impact points.

## 1. THE COVERAGE PROBLEM

### Target Value and Location Distribution

We consider—on an $x$-$y$-plane, where $x$, $y$ are Cartesian coordinates—a circularly symmetric area target whose precise position is not known because of a random locating error (see Figure 1). We assume that the probability density function (p.d.f.) $f_A(x_T, y_T)$ of the coordinates of its center is known, and that $f_A(x_T, y_T)$ is circularly symmetric with respect to the origin $O$ of the $x$-$y$-coordinate system, so that $f_A(x_T, y_T) = f[(x_T^2 + y_T^2)^{1/2}]$. The point $O$ is the expected target center location.

The area target value need not have a uniform density; for many applications it has proven useful to consider the concept of a nonuniform target value density function $w_A(u, \mathrm{v})$ as done, for example, by Eckler [5], Eckler and Burr [6], Helgert [17], McNolty [22], Schroeter [28], [29], [30], [31], and Washburn [39]. The function $w_A(u, \mathrm{v})$ need not be greater than or equal to zero, although in the usual applications it is. Our following results will apply to any value density function that is absolutely integrable over the $u$-v-plane, and circularly symmetric with respect to the center $T$ of the target, so that $w_A(u, \mathrm{v}) = w[(u^2 + \mathrm{v}^2)^{1/2}]$ within a $u$-v-coordinate system whose origin is at $T$.

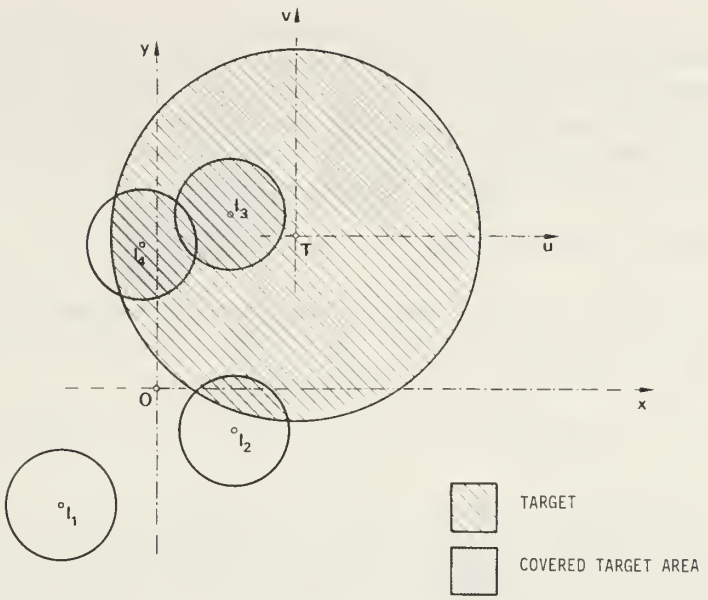



FIGURE 1. Coverage of circular target by 4 'cookie-cutter' rounds.

## P.D.F. of Impact Points of Weapons

In the following we consider a salvo of $N$ identical rounds. We assume that the weapons are aimed at the origin $O$ of the $x$-$y$-plane, the expected location of the target center. (Aiming all weapons at that point will not necessarily maximize the expected target value destroyed.) Because of random delivery errors, however, the weapons will not impact at $O$, but at points $I_j(x_j, y_j)$, $j = 1, \ldots, N$, more or less dispersed around $O$. We consider the set of impact points as a sample, randomly and independently drawn from a population with the known p.d.f. $k_A(x, y)$ which is circularly symmetric and centered at $O$, so that $k_A(x, y) = k[(x^2 + y^2)^{1/2}]$.

## Damage Function of Weapon and Coverage of Area Target

We introduce a Cartesian $\xi$-$\eta$-coordinate system whose origin is at the impact point of a weapon. We assume that the effectiveness of the weapon can be adequately described by a circularly symmetric damage function $d_A(\xi, \eta) = d[(\xi^2 + \eta^2)^{1/2}]$, with $0 \leqslant d_A(\xi, \eta) \leqslant 1$. This function has the following interpretation: The value of that element of an area target which lies at the point with the coordinates $\xi, \eta$ is reduced by the factor $1 - d_A(\xi, \eta)$. This implies that the value density of the target remaining after the bombardment by the weapon is $w_A^{(1)}(u, \mathrm{v}) = w_A(u, \mathrm{v})\ [1 - d_A(u - \xi_T, \mathrm{v} - \eta_T)]$, wherein $\xi_T, \eta_T$ are the target center coordinates in the $\xi - \eta$-coordinate system, or the weapon impact point coordinates in the $u$-v-coordinate system whose origin is at the target center. (This property also applies to any noncircularly symmetric functions $w_A$ and $d_A$, if only the axes of the $u$-v-and the $\xi$-$\eta$-coordinate systems are respectively parallel. The superscript in parentheses indicates the number of weapons considered.) Hence, the total destroyed value (or coverage) $C^{(1)}(\xi_T, \eta_T)$ is the following functional of $w_A$ and $d_A$:

$$C^{(1)}(\xi_T, \eta_T) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} w_A(u, \mathrm{v}) d_A(u - \xi_T, \mathrm{v} - \eta_T) du d\mathrm{v} \tag{1}$$

$$= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} w_A(\xi - \xi_T, \eta - \eta_T) d_A(\xi, \eta) d\xi d\eta,$$

This is the two-dimensional convolution of the two functions $w_A(u, \mathrm{v})$ and $d_A(-\xi, -\eta)$, or $w_A(-u, -\mathrm{v})$ and $d_A(\xi, \eta)$.

If both, $w_A$ and $d_A$ are the "cookie-cutter" functions for the two geometric figures $\mathscr{W}$ and $\mathscr{D}$, respectively, (i.e., when they equal one within $\mathscr{W}$ and $\mathscr{D}$, respectively, and zero without), then $C^{(1)}(\xi_T, \eta_T)$ is the intersecting area of $\mathscr{W}$ and $\mathscr{D}$. This is, of course, the usual definition of the coverage of one geometric figure by another. Using (1), one can immediately generalize this concept to more general targets and damage functions, and this definition has been applied in all papers where nonuniformly-valued area targets have been considered.

**Coverage of Area Target by Multiple Weapons**

Remember that the target center and weapon impact points coordinates in the $x$-$y$-coordinate system are $x_T, y_T, x_1, y_1, \ldots, x_N, y_N$, respectively. The coverage of the target by all the $N$ weapons then is the following function (cf. Schroeter [31], Equation (6)):

$$(2a) \quad C^{(N)}(x_T, y_T, x_1, y_1, \ldots, x_N, y_N) = K - \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} w_A(x - x_T, y - y_T) \prod_{j=1}^{N} [1 - d_A(x - x_j, y - y_j)]\,dxdy,$$

wherein $K$ is the total value of the target, $K = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} w_A(u, v)\,dudv = 2\pi\int_0^{\infty} w(t)\,t\,dt$. Define $C = C^{(N)}$ and denote the not-destroyed value $K$-$C$ by $\bar{C}$. (The complement of the destroyed value or coverage is sometimes termed "vacancy" (Ailam [1], Holst [18])). Define $\bar{d}_A(\xi, \eta) = \bar{d}[(\xi^2 + \eta^2)^{1/2}] = 1 - d_A(\xi, \eta)$. Then

$$(2b) \quad \bar{C}(x_T, y_T, x_1, y_1, \ldots, x_N, y_N) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} w_A(x - x_T, y - y_T) \prod_{j=1}^{N} \bar{d}_A(x - x_j, y - y_j)\,dxdy.$$

Considering the target center and weapon impact point coordinates as random variables, we will determine the variance of $\bar{C}$, which is the same as the variance of $C$.

**Coverage of Area Target and Probability of Killing a Point Target**

In this paper, we will not go deeper into the analysis of point targets. It is, however, interesting to relate the concept of the coverage of an area target with the probability of killing a point target, since this will enable us to generalize, in Section 2, a theorem of Robbins [25] on the relation between the moments of the not-destroyed value of the area target and the probability of killing none of several point targets. This theorem has been applied by several authors to calculating the second moment of the coverage (as done, for example, by Garwood [13]).

Consider a point target which will either survive an attack by a weapon or be killed. Define $d_A(\xi, \eta)$ to be the conditional probability of killing the point target provided it has the coordinates $\xi, \eta$ in the coordinate system whose origin is at the weapon impact point. If the point target is located randomly round the point $T(\xi_T, \eta_T)$ according to the probability density function $w_A(u, v)$, then $C^{(1)}(\xi_T, \eta_T)$ as defined by (1) is the unconditional probability of killing the point target. (This probability is unconditional with respect to the random location of the point target round $T$; it is a conditional probability with respect to the location of $T$ itself.) Evidently, $C^{(N)} = C$ as defined by (2a) equally yields the probability of the randomly located point target being killed by at least one round of the salvo. Hence, the concepts of coverage of area target and probability of kill of point target are equivalent, when $d_A$ is interpreted in the dual form, as done above, against point and area targets. Also, when $x_T, y_T, x_1, y_1, \ldots, x_N, y_N$ are considered as random variables, then the expectation $E_C$ of the destroyed value of the area target equals the totally unconditional probability of killing the point target. This equivalence is well-known (see, e.g., McNolty [22], Schroeter [28] and [31]).

I am indebted to the referee for having pointed out the following interpretation: Let $X$ denote an indicator variable for the kill of the point target, i.e., $X = 1$, whenever the point target is killed, and $X = 0$ otherwise. Again, let $w_A(u, v)$ be a probability density function. Then $C^{(N)} = C$ from (1) and (2a) equals the conditional expectation of $X$. Similarly, $E_C$ equals the unconditional expectation of $X$,

$E_X$. Since $X$ is either 0 or 1, its distribution is completely known whenever $E_X$ is. Especially, the variance of $X$, $\sigma_X^2$, equals $E_X(1 - E_X)$. Thus, nothing would remain to say on the variance of the coverage if such were the problem to be analyzed in the present paper. Instead, what we are addressing could be described, somewhat loosely, as a target consisting of an infinite number of point targets, their frequency round $T$ being proportionate to the density function $w_A(u, v)$, and $C^{(N)}$ from (1) and (2a) denoting the fraction killed. The variance of that fraction, $\sigma_C^2$, generally is less than $E_C(1 - E_C)$.

## 2. MOMENTS OF NOT-DESTROYED VALUE

### The $m$th Moment of the Not-Destroyed Value

The $m$th moment of the not-destroyed value $\bar{C}$ is

$$E_{\bar{C}^m} = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \left[ \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} w_A(x - x_T, y - y_T) \prod_{j=1}^{N} \bar{d}_A(x - x_j, y - y_j) dx dy \right]^m \tag{3a}$$
$$f_A(x_T, y_T) \prod_{j=1}^{N} k_A(x_j, y_j) dx_T dy_T dx_1 dy_1 \cdots dx_N dy_N.$$

We write the $m$th power of the inner double integral with respect to $x, y$ as a repeated $2m$-fold integral and interchange the order of the integrations. This is allowed since the integrand is an absolutely integrable function of each of its arguments. Thus, we obtain

$$E_{\bar{C}^m} = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} f_A(x_T, y_T) \prod_{i=1}^{m} w_A(\xi_i - x_T, \eta_i - y_T) \prod_{j=1}^{N} k_A(x_j, y_j) \prod_{l=1}^{m} \bar{d}_A(\xi_l - x_j, \eta_l - y_j)$$
$$dx_1 dy_1 \cdots dx_N dy_N dx_T dy_T d\xi_1 d\eta_1 \cdots d\xi_m d\eta_m.$$

The inner repeated integral with respect to $x_1, y_1, \ldots, x_N, y_N$ may now be written as the $N$th power of a double integral, so that

$$E_{\bar{C}^m} = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} f_A(x_T, y_T) \prod_{i=1}^{m} w_A(\xi_i - x_T, \eta_i - y_T) \tag{3b}$$
$$\left[ \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} k_A(\xi, \eta) \prod_{i=1}^{m} \bar{d}_A(\xi_i - \xi, \eta_i - \eta) d\xi d\eta \right]^N dx_T dy_T d\xi_1 d\eta_1 \cdots d\xi_m d\eta_m.$$

The integral of $f_A(x_T, y_T) \prod_{i=1}^{m} w_A(\xi_i - x_T, \eta_i - y_T)$ with respect to $x_T, y_T$ is completely analogous to the integral of $k_A(\xi, \eta) \prod_{i=1}^{m} \bar{d}_A(\xi_i - \xi, \eta_i - \eta)$ with respect to $\xi, \eta$. This suggests the following functional be defined:

$$\tilde{Q}_m(\xi_1, \eta_1, \ldots, \xi_m, \eta_m; k_A, \bar{d}_A) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} k_A(\xi, \eta) \prod_{i=1}^{m} \bar{d}_A(\xi_i - \xi, \eta_i - \eta) d\xi d\eta, \tag{4}$$

which implies

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f_A(x_T, y_T) \prod_{i=1}^{m} w_A(\xi_i - x_T, \eta_i - y_T) dx_T dy_T = \tilde{Q}_m(\xi_1, \eta_1, \ldots, \xi_m, \eta_m; f_A, w_A). \tag{5}$$

Using (4) and (5) we obtain for the $m$th moment of the not-destroyed value as an alternative form of (3b):

$$E_{\bar{C}^m} = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \tilde{Q}_m(\xi_1, \eta_1, \ldots, \xi_m, \eta_m; f_A, w_A) \tag{3c}$$
$$\tilde{Q}_m^N(\xi_1, \eta_1, \ldots, \xi_m, \eta_m; k_A, \bar{d}_A) d\xi_1 d\eta_1 \cdots d\xi_m d\eta_m.$$

This shows that $E_{\bar{C}^m}$ is an integral of a product, whose second factor is independent of the target value density function and the target location error p.d.f.; this factor contains all information on the weapons and on the p.d.f. of their impact points. The first factor contains all information on the target and its location error, and is independent of damage functions and impact point p.d.f.'s.

The usefulness of (3b) in contradistinction to (3a) will become apparent when we consider the second moment, at the same time taking into account that, in real situations, the number of weapons $N$ fired in a salvo generally is much greater than 2. Both, in (3a) and in (3b), the inner integrals are double integrals; the outer integrals are $(2N + 2)$-fold in (3a) and $(2m + 2)$-fold in (3b). For $N > m$, therefore, a reduction of the number of integration processes is possible by using (3b) instead of (3a).

### Interpretation of the Functional $\tilde{Q}_m$

For $m > 1$, $\tilde{Q}_m$ does not satisfy the commutative and associative laws for its argument functions, such as the ordinary convolution does. It has, however, a statistical interpretation similar to that of the convolution: Let $w_A(\xi, \eta)$ and $f_A(x, y)$ be two p.d.f.'s. Then $\tilde{Q}_m(\xi_1, \eta_1, \ldots, \xi_m, \eta_m; f_A, w_A)$ is the joint p.d.f. of the coordinates $\xi_i, \eta_i$ of $m$ points $B_i$, $i = 1, \ldots, m$, which are located randomly and independently according to the p.d.f. $w_A$ around a common center $T$ which, for its part, is drawn randomly from the $x$-$y$-plane according to the p.d.f. $f_A(x, y)$. Another interpretation results from (2b); evidently

$$\bar{C}(x_T, y_T, x_1, y_1, \ldots, x_N, y_N) = \tilde{Q}_N(x_T - x_1, y_T - y_1, \ldots, x_T - x_N, y_T - y_N; w_A, \bar{d}_A).$$

### Robbins's Theorem

An examination of (4) will show that $\tilde{Q}_m(\xi_1, \eta_1, \ldots, \xi_m, \eta_m; k_A, \bar{d}_A)$ is the probability of killing *none* of $m$ different point targets, located at the points $B_i(\xi_i, \eta_i)$, by a single round whose damage function is $d_A$ and whose impact point p.d.f. is $k_A(\xi, \eta)$. Hence, $\tilde{Q}_m^N(\xi_1, \eta_1, \ldots, \xi_m, \eta_m; k_A, \bar{d}_A)$ is the probability of killing none of these point targets by a salvo of $N$ identical and independent rounds. Interpreting, as in the section before, $\tilde{Q}_m(\xi_1, \eta_1, \ldots, \xi_m, \eta_m; f_A, w_A)$ as a joint p.d.f., we see from (3c) that the $m$th moment of the not-destroyed value of the area target, $E_{\bar{C}^m}$, equals the unconditional probability of killing *none* of $m$ randomly located point targets—which are distributed according to the p.d.f. $w_A$ round the common randomly located center $T$—with the salvo of $N$ rounds. This is the generalization of Robbins's [25] result on the moments of the coverage of a fixed geometrical figure by randomly thrown figures to the coverage of a randomly located target with arbitrary value density function by weapons with general damage functions. Evidently, these results equally apply to coverage problems where the underlying functions are not circularly symmetric.

### Expressions for the Second Moment

Because of the assumed circular symmetry of $w_A$, $f_A$, $\bar{d}_A$, and $k_A$, the integrals

$$\tilde{Q}_2(\xi_1, \eta_1, \xi_2, \eta_2; f_A, w_A) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_A(\xi, \eta)\, w_A(\xi_1 - \xi, \eta_1 - \eta)\, w_A(\xi_2 - \xi, \eta_2 - \eta)\, d\xi\, d\eta,$$

$$\tilde{Q}_2(\xi_1, \eta_1, \xi_2, \eta_2; k_A, \bar{d}_A) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} k_A(\xi, \eta)\, \bar{d}_A(\xi_1 - \xi, \eta_1 - \eta)\, \bar{d}_A(\xi_2 - \xi, \eta_2 - \eta)\, d\xi\, d\eta$$

depend only on the relative positions of the three points $O(0, 0)$, $B_1(\xi_1, \eta_1)$, and $B_2(\xi_2, \eta_2)$. Therefore we may, without loss of generality, express them in terms of the three parameters $t_1$, $t_2$, $\theta$ whose geometrical interpretation is indicated in Figure 2. A routine manipulation then yields for the second moment of the not-destroyed value:

$$E_{\bar{C}^2} = 2\pi \int_0^{\infty}\int_0^{\infty}\int_0^{2\pi} Q_2(t_1, t_2, \theta; f, w) Q_2^N(t_1, t_2, \theta; k, \bar{d})\, t_1 t_2\, d\theta\, dt_1 dt_2, \tag{6}$$

wherein

$$Q_2(t_1,\ t_2,\ \theta;\ k,\ \bar{d}) = \int_0^{\infty}\int_0^{2\pi} k(\tau)\bar{d}[(\tau^2 + t_1^2 - 2\tau t_1\cos\phi)^{1/2}]$$
$$\bar{d}\{[\tau^2 + t_2^2 - 2\tau t_2\cos(\phi - \theta)]^{1/2}\}\ \tau\, d\phi\, d\tau,$$

and the analogous relation holds for $Q_2(t_1,\ t_2,\ \theta;\ f,\ w)$. If we change from $t_1,\ t_2,\ \theta$ to $x,\ y,\ \rho$ by means of $t_1 = [(x+\rho)^2 + y^2]^{1/2}$, $t_2 = [(x-\rho)^2 + y^2]^{1/2}$, $\theta = \arctan[(x+\rho)/y] - \arctan[(x-\rho)/y]$, the Jacobian is $4\rho/(t_1 t_2)$. (The geometric interpretation of $x,\ y,\ \rho$ also is indicated in Figure 2.) Hence,

$$E_{\bar{C}^2} = 8\pi\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_0^{\infty}\hat{Q}_2(x,\ y,\ \rho;\ f_A,\ w_A)\hat{Q}_2^N(x,\ y,\ \rho;\ k_A,\ \bar{d}_A)\rho\ d\rho\, dx dy, \tag{7}$$

wherein

$$\hat{Q}_2(x,\ y,\ \rho;\ k_A,\ \bar{d}_A) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} k_A(u-x,\ v-y)\bar{d}_A(u-\rho,\ v)\bar{d}_A(u+\rho,\ v)du\, dv, \tag{8}$$

and the analogous relation holds for $\hat{Q}_2(x,\ y,\ \rho;\ f_A,\ w_A)$. Considering that $\bar{d}_A = 1 - d_A$, (8) may be transformed into

$$\hat{Q}_2(x,\ y,\ \rho;\ k_A,\ \bar{d}_A) = 1 - Q_1(t_1;\ k,\ d) - Q_1(t_2;\ k,\ d) + \hat{Q}_2(x,\ y,\ \rho;\ k_A,\ d_A), \tag{9}$$

wherein

$$Q_1(t;\ k,\ d) = \int_0^{\infty}\int_0^{2\pi} k(r)d[(t^2 + r^2 - 2tr\cos\phi)^{1/2}]r\ d\phi\, dr.$$

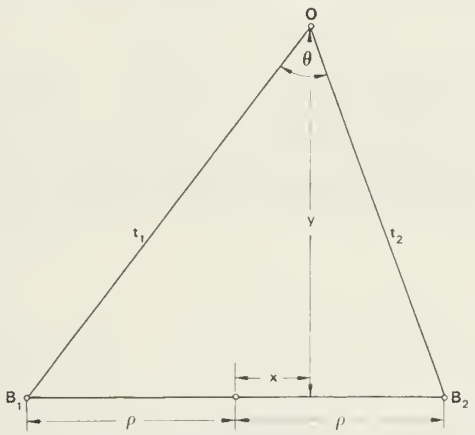



FIGURE 2. Definition of $t_1,\ t_2,\ \theta$ and $x,\ y,\ \rho$.

## 3. SECOND MOMENT IN SPECIAL CASES: GAUSSIAN AND UNIFORM FUNCTIONS

### Gaussian Damage and Impact Point Probability Density Functions

We consider the circular-Gaussian damage function $d_A(u,\ v) = \exp[-(u^2 + v^2)/(2\sigma_D^2)]$, with the shape parameter $\sigma_D$. Furthermore, we assume that the impact point p.d.f. $k_A(\cdot\ ,\ \cdot)$ is circular-normal, with the standard deviation $\sigma$. It follows immediately from the convolution property of Gaussian functions that $Q_1(t;\ k,\ d) = \sigma_D^2(\sigma^2 + \sigma_D^2)^{-1}\exp[-t^2/(2\sigma^2 + 2\sigma_D^2)]$. The function $\hat{Q}_2(x,\ y,\ \rho;\ k_A,\ d_A)$ can also be written integral-free. A routine manipulation results in

$$\hat{Q}_2(x,\ y,\ \rho;\ k_A,\ d_A) = (2\pi\sigma^2)^{-1}\exp(-\rho^2/\sigma_D^2)$$
$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\exp[-(u^2 + v^2)/\sigma_D^2]\exp\{-[(u-x)^2 + (v-y)^2]/(2\sigma^2)\}\ du\, dv.$$

We recognize the double integral as the two-dimensional convolution of two Gaussian functions, so that this integral can be expressed as a Gaussian function of $x$ and $y$. Consequently:

$$\hat{Q}_2(x, y, \rho; k_A, d_A) = \sigma_D^2(2\sigma^2 + \sigma_D^2)^{-1}\exp[-\rho^2/\sigma_D^2 - (x^2 + y^2)/(2\sigma^2 + \sigma_D^2)]. \tag{10}$$

**Gaussian Target Value and Location Probability Density Functions**

For a Gaussian target with the total value of one and the shape parameter $\sigma_Z$, we have $w_A(u, v) = (2\pi\sigma_Z^2)^{-1}\exp[-(u^2 + v^2)/(2\sigma_Z^2)]$. Let the p.d.f. of the coordinates of its center be circular-normal, with the standard deviation $\sigma_T$. Then, quite similar to the above case, $Q_1(t; f, w) = [2\pi(\sigma_Z^2 + \sigma_T^2)]^{-1}\exp[-t^2/(2\sigma_Z^2 + 2\sigma_T^2)]$ and

$$\hat{Q}_2(x, y, \rho; f_A, w_A) = [4\pi^2\sigma_Z^2(2\sigma_T^2 + \sigma_Z^2)]^{-1}\exp[-\rho^2/\sigma_Z^2 - (x^2 + y^2)/(2\sigma_T^2 + \sigma_Z^2)]. \tag{11}$$

Combining these value density and location p.d.f.'s with the Gaussian damage functions and the normal impact point p.d.f. will, for the expectation of the not-destroyed value, yield

$$E_{\bar{C}} = (\sigma_T^2 + \sigma_Z^2)^{-1}\int_0^\infty \exp\left[-\frac{t^2}{2(\sigma_T^2 + \sigma_Z^2)}\right]\left\{1 - \frac{\sigma_D^2}{\sigma^2 + \sigma_D^2}\exp\left[-\frac{t^2}{2(\sigma^2 + \sigma_D^2)}\right]\right\}^N t\, dt. \tag{12}$$

Observing (9) and inserting (10) and (11) into (7) will, for the second moment of the not-destroyed value, yield a triple integral with respect to $x, y, \rho$. Then, a change to polar coordinates by $x = r\cos\phi$, $y = r\sin\phi$ will finally result in

$$E_{\bar{C}^2} = \frac{8}{\pi\sigma_Z^2(2\sigma_T^2 + \sigma_Z^2)}\int_0^\infty\int_0^\infty \exp\left[-\frac{\rho^2}{\sigma_Z^2} - \frac{r^2}{2\sigma_T^2 + \sigma_Z^2}\right] r\rho$$
$$\int_0^{\pi/2}\left\{1 - \frac{2\sigma_D^2}{\sigma^2 + \sigma_D^2}\exp\left[-\frac{\rho^2 + r^2}{2(\sigma^2 + \sigma_D^2)}\right]\cosh\left(\frac{\rho r\cos\phi}{\sigma^2 + \sigma_D^2}\right)\right.$$
$$\left. + \frac{\sigma_D^2}{2\sigma^2 + \sigma_D^2}\exp\left[-\frac{\rho^2}{\sigma_D^2} - \frac{r^2}{2\sigma^2 + \sigma_D^2}\right]\right\}^N d\phi\, dr d\rho. \tag{13}$$

The integrals (12) and (13) can be solved in closed form, by developing the $N$th powers and integrating the sums term by term. Practically, this is of little interest, since the complexity of the integrands is prohibitive for large $N$. (12) and (13), however, can be computed easily by numerical quadrature for arbitrary values of $N$.

**Coverage of Fixed Circle by Randomly Thrown Circles**

In this section we apply our main result of Section 2 to generalizing results on a circularly symmetric coverage problem considered by Garwood [13] and Santaló [27]. These authors calculated the second moment of the coverage of a fixed circle (radius $R_T$) by $N$ identical, randomly thrown circles (radius $a$) when the centers of the latter are distributed uniformly over a circle with radius $R_D$. The centers of the two circles with the radii $R_T$ and $R_D$ coincide.

The damage function $d(\tau)$, in this case, is 1 if $\tau \leqslant a$, and 0 else. We denote by $k_A(u, v)$ the uniform p.d.f. over the circle with the radius $R_D$. Then $\pi R_D^2 Q_1(t; k, d)$ is the intersecting area $S_1(t; R_D, a)$ of the two circles with the radii $a$ and $R_D$ whose centers are the distance $t$ apart. Furthermore, $\pi R_D^2 Q_2(t_1, t_2, \theta; k, d)$ is the intersecting area of the three circles with the radii $R_D$, $a$, $a$, and

the centers at the corners $O$, $B_1$, $B_2$, respectively, of the triangle indicated in Fig. 2. We denote this intersecting area of the three circles mentioned by $S_2(t_1, t_2, \theta; R_D, a)$ and, using (9), obtain

$$Q_2(t_1, t_2, \theta; k, \bar{d}) = 1 - (\pi R_D^2)^{-1}[S_1(t_1; R_D, a) + S_1(t_2; R_D, a) - S_2(t_1, t_2, \theta; R_D, a)]. \tag{14}$$

Next we consider the target whose value density is one within a circle of radius $R_T$, and zero without. We assume that this target is centered, with the probability one, at the point $O(0, 0)$. In this case $Q_2(t_1, t_2, \theta; f, w) = 1$ if $t_1 \leqslant R_T$ and $t_2 \leqslant R_T$, and 0 otherwise. Hence (6) yields

$$E_{\bar{C}^2} = 2\pi \int_0^{R_T} \int_0^{R_T} \int_0^{2\pi} Q_2^N(t_1, t_2, \theta; k, \bar{d}) t_1 t_2 d\theta dt_1 dt_2.$$

Combining this with (14) results in the second moment for that proportion of the fixed circle which is not covered by the $N$ randomly thrown circles of radius $a$:

$$E_{\bar{C}^2} = 2\pi \int_0^{R_T} \int_0^{R_T} \int_0^{2\pi} \{1 - (\pi R_D^2)^{-1}[S_1(t_1; R_D, a) + S_1(t_2; R_D, a) - S_2(t_1, t_2, \theta; R_D, a)]\}^N t_1 t_2 d\theta dt_1 dt_2. \tag{15}$$

This is the generalization to arbitrary relations of $R_T$, $R_D$, and $a$, of results of Garwood and Santaló, who considered only the case $R_D - a \geqslant R_T$. In the latter case, (15) may be reduced to a one-fold integral; a short manipulation results in

$$E_{\bar{C}^2} = 2\pi \int_0^{2R_T} S_1(z; R_T, R_T)\{1 - (\pi R_D^2)^{-1}[2\pi a^2 - S_1(z; a, a)]\}^N z dz,$$

which is an alternative form of Garwood's Equation (43) and Santaló's Equation (5.3).

## 4. NUMERICAL RESULTS FOR GAUSSIAN FUNCTIONS

Equations (12) and (13) were evaluated by numerical integration, and the standard deviation $\sigma_C$ of the destroyed value was determined by means of the relation $\sigma_C^2 = \sigma_{\bar{C}}^2 = E_{\bar{C}^2} - (E_{\bar{C}})^2$. In all cases we assumed $\sigma_Z = 1$. Figure 3-6 show in the $E_C$–$\sigma_C$-plane the two families of parameter curves $\sigma_T$ = const. and $N$ = const. for $\sigma_D = 0.1$ and 0.2, $\sigma = 1$ and 3. Many practically important cases fall within the range of these parameters.

### Discussion of Results

The figures indicate that, for $\sigma_T \leqslant 1$, the standard deviation of the coverage is so small that the one-sided Tchebyshev inequality $P(E_C - \alpha\sigma_C \leqslant C) \geqslant \alpha^2/(1 + \alpha^2)$ may be well applied to deriving a confidence level for the minimum destroyed value. Take, e.g., $\sigma_D = 0.1$, $\sigma = 1$, $\sigma_T = 0.5$, $N = 250$ rounds. It can be seen in Figure 3 that $E_C \cong 57\%$, $\sigma_C \cong 5\%$, so that the coverage is with a probability of at least 90% greater than 42%. For $\sigma_T >> 1$, however, Tchebyshev's inequality will be of little value for our problem.

We see that, for increasing $\sigma_T$ and $N$, the standard deviation of the coverage $\sigma_C$ approaches its theoretical upper bound on the circle with the equation $\sigma_C^2 = E_C(1 - E_C)$. This circle is drawn as a dashed line. This property implies that the probability mass concentrates more and more near $C = 0$ and $C = 1$.

### Numerical Method

The integrals (12) and (13) were evaluated by a Gauss-Legendre quadrature. For (13), generally $24^3$ supporting points were used. For special values of the parameters, $32^3$ supporting points were chosen. With a Fortran program, on a CDC 6500 computer approximately three seconds of central processor time was required to determine one value of $\sigma_C$ with $24^3$ supporting points.

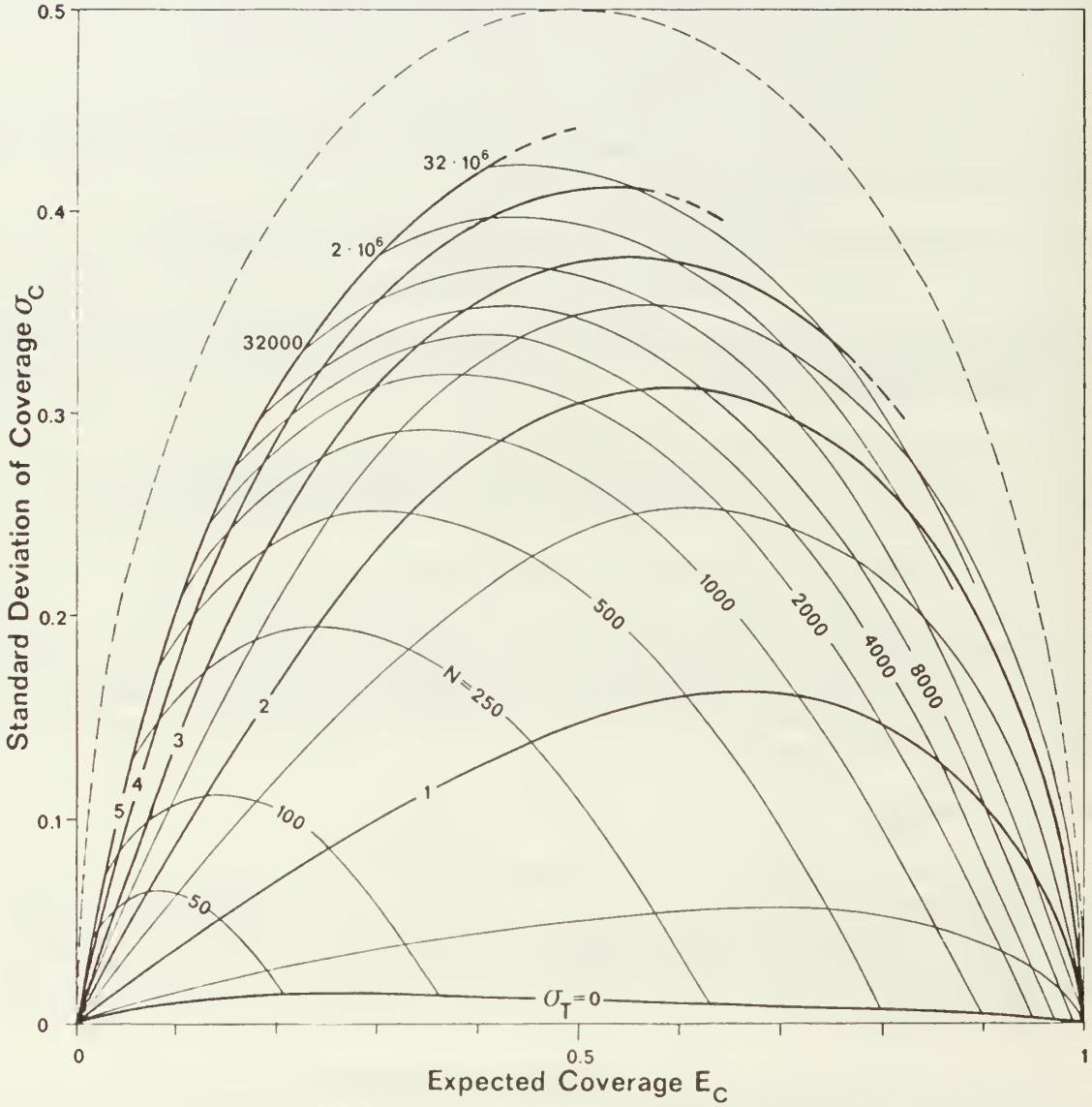



FIGURE 3. Standard deviation of coverage vs. expected coverage. $\sigma_D = 0.1$, $\sigma = 1$.

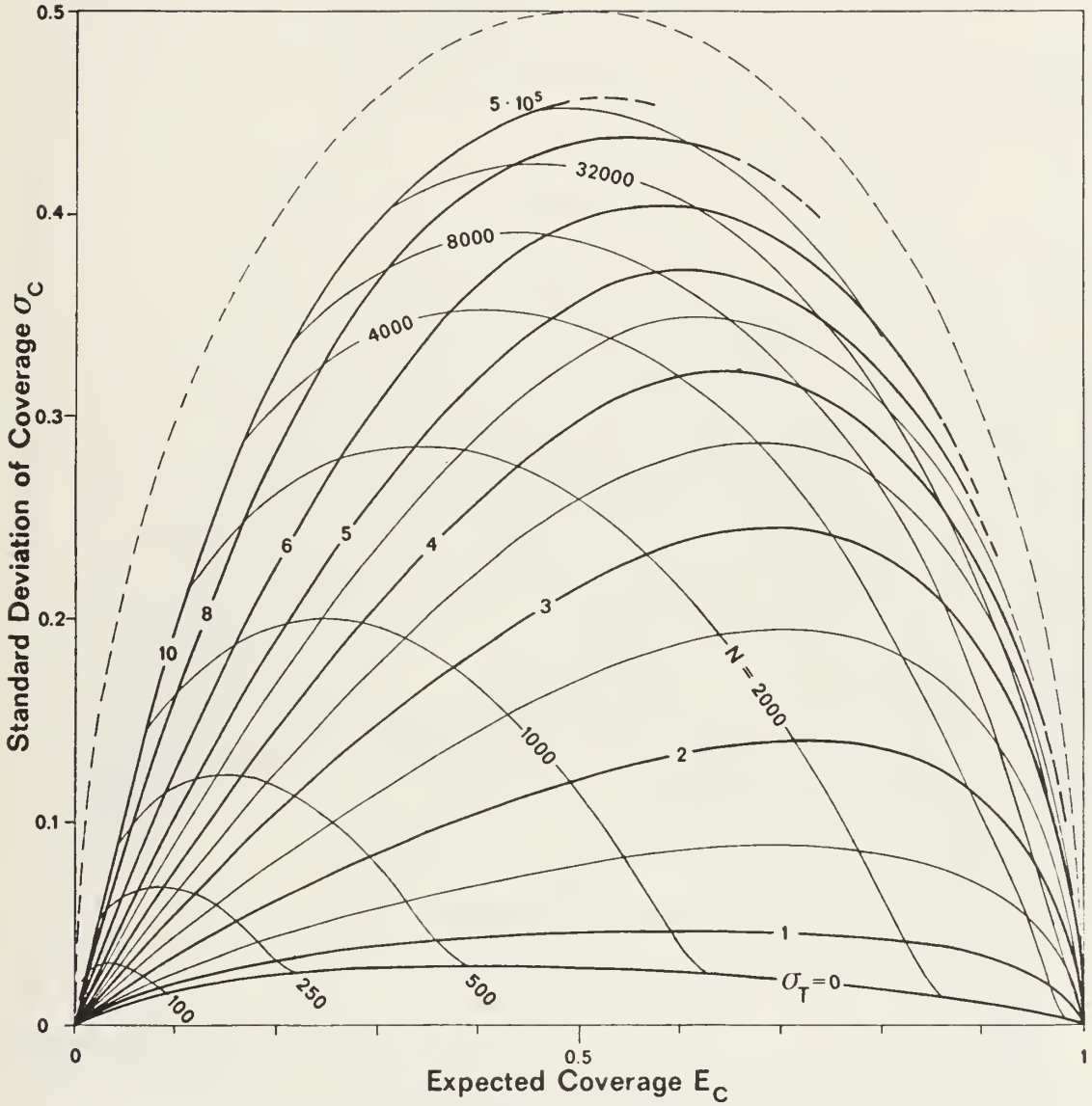



FIGURE 4. Standard deviation of coverage vs. expected coverage. $\sigma_D = 0.1$, $\sigma = 3$.

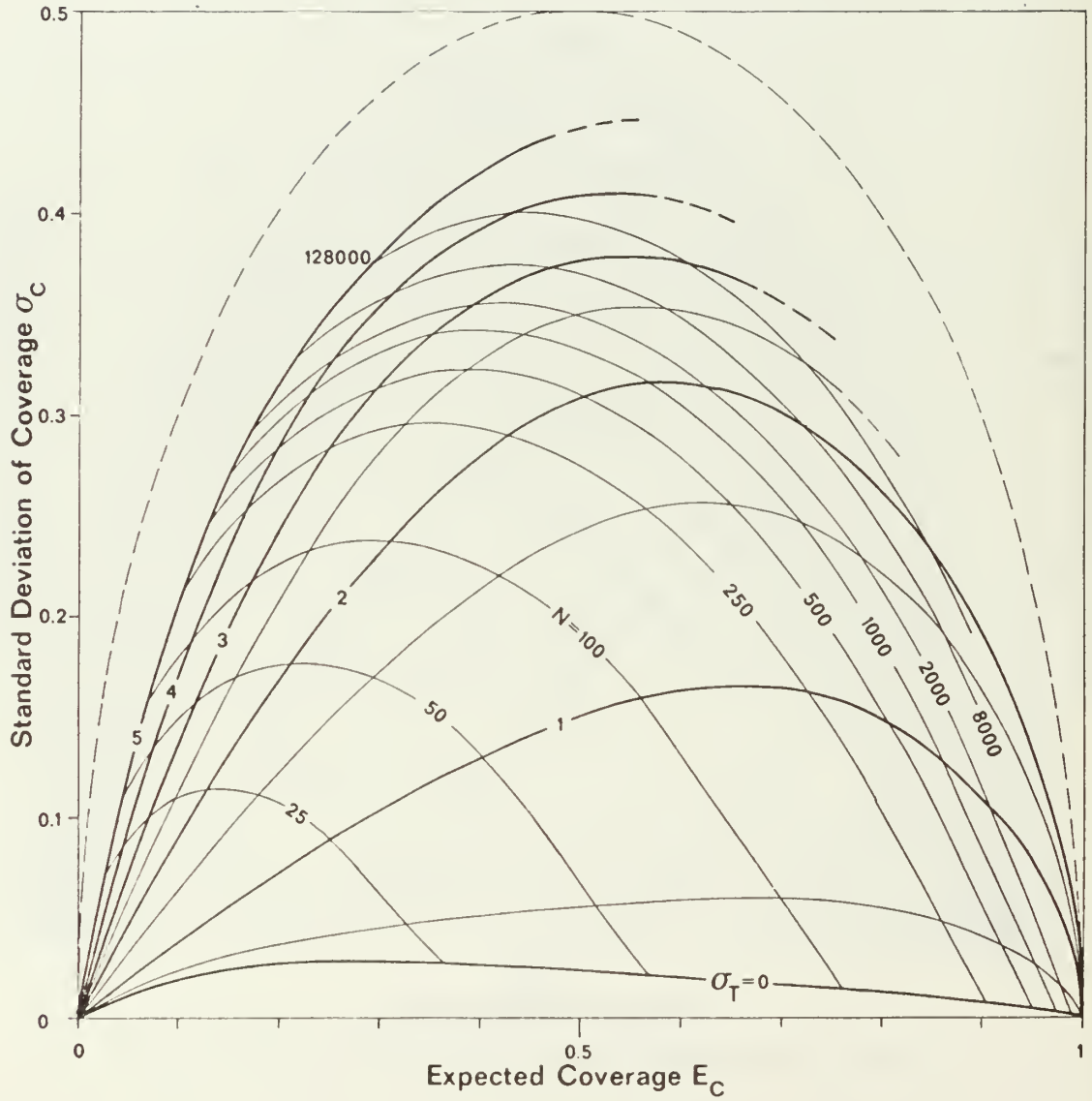



FIGURE 5. Standard deviation of coverage vs. expected coverage. $\sigma_D = 0.2$, $\sigma = 1$.

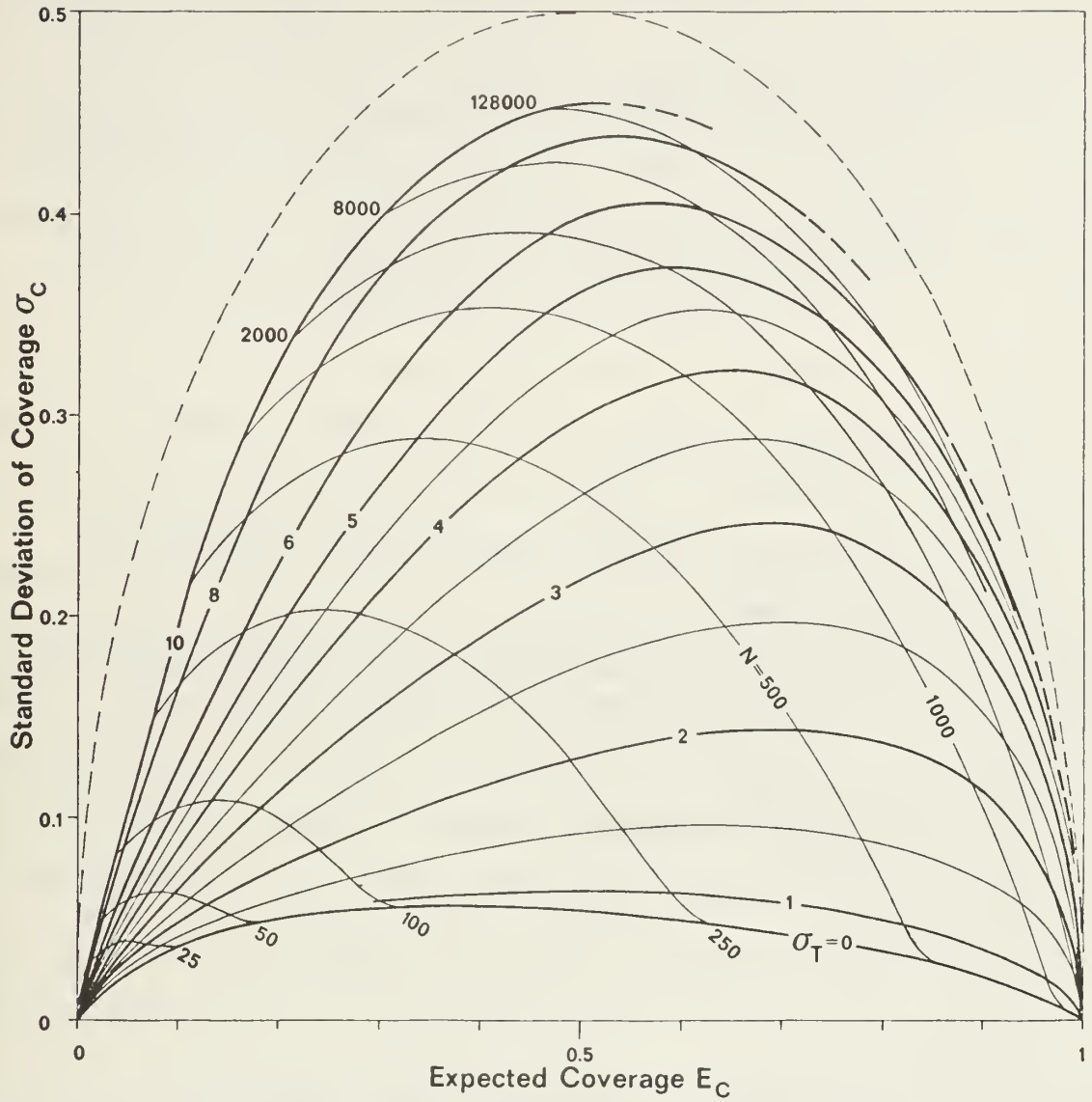



FIGURE 6. Standard deviation of coverage vs. expected coverage. $\sigma_D = 0.2$, $\sigma = 3$.

## ACKNOWLEDGMENT

This work was done for the Ministry of Defense of the Federal Republic of Germany.

## REFERENCES

[1] Ailam, G., "Moments of Coverage and Coverage Spaces," Journal of Applied Probability, *3*, 550-555 (1966).

[2] Bronowski, J. and J. Neyman, "The Variance of the Measure of a Two-Dimensional Random Set," Annals of Mathematical Statistics, *16*, 330-341 (1945).

[3] Cooke, P.J., "Bounds for Coverage Probabilities with Applications to Sequential Coverage Problems," see [1], *11*, 281-293 (1974).

[4] Domb, C., "The Problem of Random Intervals on a Line," Proceedings of the Cambridge Philosophical Society, *43*, 329-341 (1947).

[5] Eckler, A.R., "A Survey of Coverage Problems Associated with Point and Area Targets," Technometrics, *11*, 561-589 (1969).
[6] Eckler, A.R. and S.A. Burr, *Mathematical Models of Target Coverage and Missile Allocation*, Military Operations Research Society, Alexandria, VA (1972).
[7] Edens, E., "Random Covering of a Circle," Indagationes Mathematicae, *37*, 373-384 (1975).
[8] Flatto, L., "A Limit Theorem for Random Coverings of a Circle," Israel Journal of Mathematics, *15*, 167-184 (1973).
[9] Flatto, L. and A.G. Konheim, "The Random Division of an Interval and the Random Covering of a Circle," SIAM Review, *4*, 211-222 (1962).
[10] Flatto, L. and D.J. Newman, "Random Coverings," Acta Mathematica, *138*, 241-264 (1977).
[11] Fraser, D.A.S., "Generalized Hit Probabilities with a Gaussian Target," see [2], *22*, 248-255 (1951).
[12] Fraser, D.A.S., "Generalized Hit Probabilities with a Gaussian Target II," see [2], *24*, 288-294 (1953).
[13] Garwood, F., "The Variance of the Overlap of Geometrical Figures with Reference to a Bombing Problem," Biometrika, *34*, 1-17 (1947).
[14] Gilbert, E.N., "The Probability of Covering a Sphere with $N$ Circular Caps," Biometrika, *52*, 323-330 (1965).
[15] Greenberg, I., "The Moments of Coverage of a Linear Set," see [1], *17*, 865-868 (1980).
[16] Guenther, W.C. and P.J. Terragno, "A Review of the Literature on a Class of Coverage Problems," see [2], *35*, 232-260 (1964).
[17] Helgert, H.J., "On the Computation of Hit Probability," Operations Research, *19*, 668-684 (1971).
[18] Holst, L., "On Multiple Covering of a Circle with Random Arcs," see [1], *17*, 284-290 (1980).
[19] Holst, L., "On the Lengths of the Pieces of a Stick Broken at Random," see [1], *17*, 623-634 (1980).
[20] Kaplan, N., "Two Applications of a Poisson Approximation for Dependent Events," Annals of Probability, *5*, 787-794 (1977).
[21] Kaplan, N., "A Limit Theorem for Random Coverings of a Circle which Do Not Quite Cover," see [1], *15*, 443-446 (1978).
[22] McNolty, F., "Expected Coverage for Targets of Non-Uniform Density," Operations Research, *16*, 1027-1040 (1968).
[23] Miles, R.E., "The Asymptotic Values of Certain Coverage Probabilities," Biometrika, *56*, 661-680 (1969).
[24] Moran, P.A.P. and S. Fazekas De St. Groth, "Random Circles on a Sphere," Biometrika *49*, 389-396 (1962).
[25] Robbins, H.E., "On the Measure of a Random Set," see [2], *15*, 70-74 (1944).
[26] Robbins, H.E., "On the Measure of a Random Set II," Annals of Mathematical Statistics, *16*, 342-347 (1945).
[27] Santaló, L.A., "On the First Two Moments of the Measure of a Random Set," see [2], *18*, 37-49 (1947).
[28] Schroeter, G., "Probability of Kill and Expected Destroyed Value if the Underlying Distributions Are Rotationally Symmetric," Operations Research, *24*, 586-591 (1976).
[29] Schroeter, G., "The Application of a Hankel Transformation to a Class of Circularly Symmetric Coverage Problems," Operations Research, *25*, 675-687 (1977).
[30] Schroeter, G., "The Application of the Hankel Transformation to Spherically Symmetric Coverage Problems," Journal of Applied Probability, *17*, 1121-1126 (1980).
[31] Schroeter, G., "Expected Coverage of a Randomly Located Target by Multiple Independent Salvos," Operations Research, *28*, 1299-1318 (1980).
[32] Shepp, L.A., "Covering the Circle with Random Arcs," Israel Journal of Mathematics, *11*, 328-345 (1972).
[33] Siegel, A.F., "Random Space Filling and Moments of Coverage in Geometrical Probability," see [1], *15*, 340-355 (1978).

[34] Siegel, A.F., "Random Arcs on the Circle," see [1], *15*, 774-789 (1978).
[35] Solomon, H., "Distribution of the Measure of a Random Two-Dimensional Set," Annals of Mathematical Statistics, *24*, 650-656 (1953).
[36] Steutel, F.W., "Random Division of an Interval," Statistica Neerlandica, *21*, 231-244 (1967).
[37] Stevens, W.L., "Solution to a Geometrical Problem in Probability," Annals of Eugenics, London, *9*, 315-320 (1939).
[38] Votaw, Jr., D.F., "The Probability Distribution of the Measure of a Random Linear Set," Annals of Mathematical Statistics, *17*, 240-244 (1946).
[39] Washburn, A.R., "Upper and Lower Bounds for a Pattern Bombing Problem," Naval Research Logistics Quarterly, *21*, 705-713 (1974).

# REFORMULATING ZERO-SUM GAMES WITH MULTIPLE GOALS

Edward L. Hannan

*Office of Health Systems Management*
*New York State Department of Health*
*Albany, New York*

ABSTRACT

This paper investigates the two-person zero-sum multiple payoff game in which the objective is to minimize a player's total underachievement from a fixed set of goals. It is demonstrated that a previous formulation of this problem can be substantially simplified.

## 1. INTRODUCTION

In a previous paper in NRLQ, Cook [1] discussed the formulation and solution of a two-person, zero-sum game with multiple payoffs in which the objective is to minimize the total underachievement from a fixed set of goals for each of the objectives.

The resulting formulation is

$$\min_{x \in \chi, \beta} \beta \tag{1}$$

subject to

$$\sum_{i=1}^{m} x_i A_{ij}(r) - \beta \leqslant 0 \; \forall \mathrm{j,r}; x \in \chi$$

where player 1 selects strategy $i$ with probability $x_i$,

$$X = (x_1, x_2, \ldots, x_m)$$

$$\chi = \left\{ X = (x_1, \ldots, x_m) \qquad \sum_{i=1}^{m} x_i = 1, \; x_i \geqslant 0 \; \forall \mathrm{i} \right\}$$

$$A(r) = \sum_{k=1}^{K} \hat{A}^k \alpha_r^k$$

$$\hat{A}^k = (\hat{a}_{ij}^k) = (g^k - a_{ij}^k)$$

$K$ is the number of objectives

$j$ is the index for player 2's strategies

$g^k$ is the goal for the $k$th objective

$a_{ij}^k$ is the payoff for the $k$th objective corresponding to player 1 choosing strategy $i$ and player 2 choosing strategy $j$

$A_{ij}(r)$ is the $ij$th component of $A(r)$

$\alpha^k$ is the set of extreme points of $\psi$

$$\psi = \left\{\alpha = (\alpha^1, \ldots, \alpha^k) \quad 0 \leqslant \alpha^k \leqslant w^k\right\}.$$

$r = 1, \ldots, R$ is the index on the extreme points of $\psi$.

$w^k$ = the relative weight placed on the $k$th goal.

This is a linear programming problem. After calculating the $A(r)$ matrices, the problem involves the minimization of a linear function subject to $m(2^K - 1) + 1$ constraints where $m$ is the number of strategies available to player 1 and $K$ is the number of multiple payoffs. Also, the calculation of the $2^K$ $A(r)$ matrices prior to solving the problem can be quite time-consuming.

The intent of this paper is to demonstrate that the same problem can be solved by minimizing a linear function subject to $n(K + 1) + 1$ constraints, where $n$ is the number of strategies available to player 2. Also, only $K$ matrices must be calculated prior to the solution of the problem rather than $2^K$ matrices.

## 2. A NEW FORMULATION

Letting $a_{ij}^k$ represent the $k$th payoff to player 1 if player 1 chooses strategy $i$ with probability $x_i$, $i = 1, \ldots m$, and player 2 chooses strategy $j$, $j = 1, \ldots, n$, it follows that $\sum_{i=1}^{m} a_{ij}^k x_i$ is the $k$th payoff to player 1 for this combination of strategies.

Thus,

$$d_j^k = \max\left[0, g^k - \sum_{i=1}^{m} a_{ij}^k x_i\right]$$

is the underachievement of goal $k$ ($g^k$) corresponding to this strategy combination.

But $d_j^k = \max\left[0, \sum_{i=1}^{m} (g^k - a_{ij}^k) x_i\right]$

since $g^k$ is constant and $\sum_{i=1}^{m} x_i = 1$.

Since each of these $d^k$ values are weighted by corresponding $w^k$ values, player 1 selects $x_i$ $\left[x_i \geqslant 0, \sum_{i=1}^{m} x_i = 1\right]$ which will yield

$$\min_{x_i}\left\{\max\left[\sum_{k=1}^{K} w^k d_1^k, \ldots, \sum_{k=1}^{K} w^k d_n^k\right]\right\}$$

where $K$ is the number of multiple payoffs.

Hence, the problem can be expressed as

(2) $\min v$

subject to

$$v \geqslant \sum_{k=1}^{K} w^k d_j^k \quad \forall j$$

$$d_j^k = \max \left\{ \sum_{i=1}^{m} (g^k - a_{ij}^k) x_i', \ 0 \right\} \forall j, \ \forall k$$

$$\sum_{i=1}^{m} x_i' = 1$$

$$x_i' \geqslant 0 \quad \forall i$$

where v is the value of the game.

However, this problem may be expressed as a linear programming problem.

THEOREM 1: Formulation (1) is equivalent to the linear programming problem

(3) $\min v$

subject to

$$v \geqslant \sum_{k=1}^{K} w^k f_j^k \quad \forall j$$

$$f_j^k \geqslant \sum_{i=1}^{m} (g^k - a_{ij}^k) \, x_i \quad \forall j, \ \forall k$$

$$\sum_{i=1}^{m} x_i = 1$$

$$x_j \geqslant 0 \ \forall j, \ f_j^k \geqslant 0 \ \forall j, \ \forall k.$$

PROOF: Suppose $v_1$ is the optimal solution to (1) and $v_2$ is the optimal solution to (2). We want to show that $v_1 = v_2$. Clearly, $v_2 \geqslant v_1$ since the $w^k$ are positive and $f_j^k \geqslant d_j^k$ for all $j$ and $k$.

Thus, it must be shown that $v_2 \leqslant v_1$. Suppose $v_2 > v_1$. Then $v_1 = \sum_{k=1}^{K} w^k d_j^k$ for some $j$, say $j = j_1$, and some set of $x_i'$ values, say $x_i' = x_{i_0}' \ \forall i$. Letting $x_{i_0} = x_{i_0}'$ for all $i$ values and $f_{j_1}^k = d_{j_1}^k$ for all $k$ values, we obtain a feasible solution for problem (2) which is equal to $v_1$ and less than $v_2$. This is a contradiction to the original assumption.

The following theorem demonstrates that formulation (3) is equivalent to Cook's formulation (1).

THEOREM 2: Formulations (1) and (3) are equivalent.

PROOF: Cook's formulation is

$$\min_x \max_y \sum_k w^k d^k(x, y)$$

where $x$ and $y$ range over $m$ and $n$ dimensional unit simplices, and

$$d^k(x, y) = \max \left\{ 0, \ \sum_i \sum_j x_i \, (g^k - a_{ij}^k) y_j \right\}.$$

Formulation (3) is

$$\min_x \max_{j=1,2,\ldots,n} \sum_k w^k d_j^k(x)$$

where

$$d_j^k(x) = \max\left\{0, \sum_i x_i\ (g^k - a_{ij}^k)\right\}$$

Clearly, $d^k(x, y) \leqslant \sum_j d_j^k(x) y_j$ for all $k$ for any $x$, $y$.

Also, since $w^k \geqslant 0$ for all $k$,

$$\sum_k w^k d^k\ (x,y) \leqslant \sum_j \left(\sum_k w^k d_j^k(x)\right) y_j$$

and

$$\max_y \sum_k w^k d^k(x, y) \leqslant \max_y \sum_j \left(\sum_k w^k d_j^k(x)\right) y_j$$
$$= \max_{j=1,2,\ldots,n} \sum_k w^k d_j^k(x)$$

But, if $e^j$ is the extreme point of the $y$-simplex whose $j$-th coordinate is 1,

$$d^k(x, e^j) = d_j^k(x).$$

Thus,

$$\max_j \sum_k w^k d_j^k(x) = \max_{j=1,2,\ldots,n} \sum_k w^k d^k\ (x, e^j).$$

Also,

$$\max_j \sum_k w^k d^k(x, e^j) \leqslant \max_y \sum_k w^k d^k(x, y)$$

since the sums considered on the left are a subset of those considered on the right.

Consequently,

$$\max_y \sum_k w^k d^k(x, y) = \max_j \sum_k w^k d_j^k(x)$$

and the two formulations are equivalent.

## 3. EXAMPLES

EXAMPLE 1: Consider the multigoal game solved by Cook [1] with three different payoffs where $g^1 = 1$, $g^2 = 5$, $g^3 = 4$ and corresponding payoff matrices given by

$$A^1 = \begin{bmatrix} 1 & -6 \\ 2 & 7 \end{bmatrix} \quad A^2 = \begin{bmatrix} -10 & 7 \\ 3 & -4 \end{bmatrix} \quad A^3 = \begin{bmatrix} 2 & 4 \\ -6 & -1 \end{bmatrix}$$

and weights $w^1 = 2$, $w^2 = .5$ and $w^3 = 1$. The problem may be formulated as

$$
\begin{array}{lll}
\min & \mathrm{v} & \\
\text{subject to} & & \\
 & \mathrm{v} \geqslant 2f_1^1 + .5f_1^2 + f_1^3 & \\
 & \mathrm{v} \geqslant 2f_2^1 + .5f_2^2 + f_2^3 & \\
-x_2 & \leqslant f_1^1 & \\
15x_1 + 2x_2 & \leqslant f_1^2 & \\
2x_1 + 10x_2 & \leqslant f_1^3 & \\
7x_1 + 8x_2 & \leqslant f_2^1 & \\
-2x_1 + 9x_2 & \leqslant f_2^2 & \\
5x_2 & \leqslant f_2^3 &
\end{array}
$$

$$\sum_{i=1}^{2} x_i = 1$$

$$x_i \geqslant 0 \;\forall \mathrm{i},\; \mathrm{f_i^k} \geqslant 0 \;\forall \mathrm{j},\; \forall \mathrm{k}.$$

The solution to this problem is $x_1 = 1$, $x_2 = 0$, $\mathrm{v} = 14$, $y_1 = 0$ $y_2 = 1$ as given by Cook. Note that in this formulation only three matrices (the $((g^k - a_{ij}^k))$, $k = 1, 2, 3$) need be calculated in advance. The eight extreme points $\alpha_r$, $r = 1, \ldots, 8$ and the eight matrices $A(r)$, $r = 1, \ldots, 8$ are not required. Also, there are 9 constraints instead of the 15 required by Cook.

EXAMPLE 2: In this example provided also by Cook, the three goals are $g^1 = 4$, $g^2 = 1$, $g^3 = 2$ with corresponding weights $w^1 = 1$, $w^2 = 2$, $w^3 = 2.5$. The payoff matrices are

$$A^1 = \begin{bmatrix} 2 & 5 & 1 \\ -1 & -2 & 6 \\ 0 & 3 & -1 \end{bmatrix}, \quad A^2 = \begin{bmatrix} -3 & 7 & 2 \\ 0 & -2 & 0 \\ 3 & -1 & 6 \end{bmatrix}, \quad A^3 = \begin{bmatrix} 8 & -2 & 3 \\ -5 & 6 & 0 \\ -3 & 1 & 6 \end{bmatrix}$$

Using the formulation given in (2), the solution to this problem is $\mathrm{v} = 6.32456$, $x_1 = .49123$, $x_2 = .20175$, $x_3 = .30702$. This solution differs from that given by Cook ($\mathrm{v} = 5.8148$, $x_1 = .636024$, $x_2 = .157764$, $x_3 = .206211$), but this is the result of some errors in the $A(r)$ matrices calculated by Cook. The corrections in these matrices which need to be made are $A_{31}(2) = 4$, $A_{33}(4) = -10$, $A_{31}(5) = 0$, $A_{31}(6) = 16.5$, $A_{33}(6) = -5$, $A_{33}(7) = 4$, $A_{31}(8) = 12.5$, and $A_{33}(8) = 9$. The errors were apparently the results of calculating $\hat{a}_{11}^1$ (the 1-1 element of $\hat{A}_1$) to be $-4$ instead of 4 and calculating $\hat{a}_{31}^3$ to be 4 instead of $-4$. When these corrections are made, Cook's formulation also produces the solution given above.

The formulation in (2) has 13 constraints and the previous formulation has 22 constraints.

## 3. SUMMARY

The formulation of the two-person zero-sum multiple payoff game given in (2) is an improvement over the previous formulation in that the number of constraints increases proiportionately with the number of strategies available to player 2 and proportionately with the number of payoffs as opposed to increasing proportionately with the number of strategies available to player 1 and exponentially with the number of payoffs. In general, the new formulation has $n(K + 1) + 1$ constraints instead of $m(2^K - 1) + 1$ constraints.

For example, with 5 payoffs present, the new formulation has 31, 25, and 19 constraints when there are 5, 4, and 3 strategies, respectively, available to each player. In the previous formulation there are 156, 125, and 94 constraints, respectively.

Finally, it should be noted that there is no appropriate formulation corresponding to (2) for the payoff to player 2 since the formulation leads to unbounded values for the $f_i^k$ and the objective function is consequently unbounded.

## ACKNOWLEDGMENT

I would like to thank an anonymous referee for demonstrating the equivalence provided in Theorem 2.

## REFERENCE

[1] Cook, W.D., "Zero-Sum Games with Multiple Goals," Naval Research Logistics Quarterly, *23*, 615-622 (1976).

# ON A GENERALIZED BIN-PACKING PROBLEM

R. T. Lewis

*R.C.A. Corporation*
*Princeton, New Jersey*

R. G. Parker

*School of Industrial and Systems Engineering*
*Georgia Institute of Technology*
*Atanta, Georgia*

**ABSTRACT**

In this paper we address a bin-packing problem which possesses a variety of modifications of the classic theme. Among these are bin-dependent chip weights, bin costs, and bin-dependent penalties for unused capacity. Lagrangian relaxations are employed in the context of a branch-and-bound framework in order to solve the problem after which substantial computational experience is provided.

## 1. INTRODUCTION

### 1.1 Problem Statement

The classic, one-dimensional bin packing problem seeks the minimum number of bins (of finite capacity) required to pack a set of chips possessing given, bin-independent displacements (weight, length, volume, etc.). In this paper, we consider a generalization of this theme which can be made precise by model (GBPP1) below:

(1) (GBPP1): Minimize: $\sum_{j=1}^{m} (c_j y_j + d_j u_j)$

(2) Subject to: $\sum_{i=1}^{n} t_{ij} x_{ij} + u_j = w_j y_j,\ 1 \leqslant j \leqslant m$

(3) $\sum_{j=1}^{m} x_{ij} = 1,\ 1 \leqslant i \leqslant n$

(4) $x_{ij},\ y_j$ binary

where: $w_j$ is a strictly positive, integer capacity of bin $j$,
$t_{ij}$ is the deplacement of chip $i$ if assigned to bin $j$,
$u_j$ is the unused capacity of bin $j$,
$c_j$ is a positive, integer cost of bin $j$, and
$d_j$ is a positive, integer cost of unused capacity of bin $j$.

The binary variables $x_{ij}$ and $y_j$ relate to the assignment of chip $i$ to bin $j$ and the use of bin $j$, respectively. We assume the problem dimension to be $n$(chips) by $m$(bins) where $n$ and $m$ are such that at least one feasible pack is possible. By "feasible pack" we mean any solution satisfying (2)-(4). Additionally, we assume the unused capacity cost (and the bin cost) to be realized only if a bin is used.

While not considered in the research presented here, it is worth remarking that it may be of interest to constrain (GBPP1) further by assuming the existence of various packing restrictions. To illustrate, suppose chips represent drums of chemicals where two chemical types which otherwise could be stored together, may possess properties making such storage undesirable. Letting $R_j$ be the set of (possibly empty) pairwise packing restrictions for bin $j$, we could append (GBPP1) by constraints of the form

$$x_{kj} + x_{hj} \leqslant 1, \quad (k, h) \in R_j, \quad 1 \leqslant j \leqslant m. \tag{5}$$

Indeed, constraints of this nature have been considered in other work by the first author [22]; however, in what follows, we have assumed $R_j = \emptyset$ for all $j$.

It is easy to see that the classic bin-packing problem is but a special case of (GBPP1). In particular, we can set $c_j = 1$, $d_j = 0$, $w_j = w$ for $1 \leqslant j \leqslant m$ and let $t_{ij} = t_i$ for $1 \leqslant i \leqslant n$, $1 \leqslant j \leqslant m$. Equivalently, letting $c_j = 0$ and $d_j = 1$ for $1 \leqslant j \leqslant m$ with $t_{ij}$, and $w_j$ as just specified achieves the same result.

While bin-dependent chip displacement and differing bin capacities provide rather natural generalizations of the classic problem, the cost penalty for unused capacity is less so. In addition, this cost penalty for unused capacity can clearly create problems where a solution using more costly bins will have a smaller total cost than a solution where less costly bins are used. To see this consider the following small illustration:

| $j$: | 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|---|
| $c_j$: | $c_1$ | $4/5c_1$ | $3/5c_1$ | $2/5c_1$ | $1/5c_1$ |
| $d_j$: | $d_1$ | $5/4d_1$ | $5/3d_1$ | $5/2d_1$ | $5d_1$ |
| $w_j$: | $c_1$ | $c_2$ | $c_3$ | $c_4$ | $c_5$ |

$t_{ij}$

| $i$ \ $j$ | 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|---|
| 1 | $c_1$ | $1/3c_2$ | $\infty$ | $\infty$ | $\infty$ |
| 2 | $1/3c_1$ | $c_2$ | $\infty$ | $\infty$ | $\infty$ |
| 3 | $\infty$ | $1/3c_2$ | $c_3$ | $\infty$ | $\infty$ |
| 4 | $\infty$ | $1/3c_2$ | $\infty$ | $c_4$ | $\infty$ |
| 5 | $1/3c_1$ | $\infty$ | $\infty$ | $\infty$ | $c_5$ |

Clearly, a pack requiring the fewest bins is achieved with $x_{21} = x_{51} = x_{12} = x_{32} = x_{42} = 1$ and $x_{ij} = 0$, elsewhere. The cost or value of this solution is $c_1 + c_2 + c_1d_1/3 = \frac{9c_1}{5} + c_1d_1/3$. Consider now the solution $x_{ij} = 1$ for all $i = j$ and $x_{ij} = 0$ elsewhere, having value $\sum_{j=1}^{5} c_j = 3c_1$. If $0 \leqslant d_1 < \frac{18}{5}$ the first solution costs less than the second. Alternately, if $d_1 > \frac{18}{5}$, the latter solution is better.

A central ingredient (although not the only one) in differentiating between least cost and classic, least cardinality packs is the assumed penalty or charge for unused bin capacity relative to bins used. As one might expect, this implicit constraint on free resource disposal creates a degree of difficulty which empirically speaking, can make for exceedingly difficult problems. Note that we use the "empirical" qualifier here to distinguish the difficulty of (GBPP1) from that of the classic bin-packing problem since otherwise the notion of difficulty or complexity is indistinguishable due to the NP-completeness of the classic problem.

It is interesting to imagine physical systems where (GBPP1) applies. One example arises in the case of so-called parallel-processor scheduling [1, 3]. Here, we seek a schedule-creating assignment of single operation jobs to a set of processors which minimizes the completion time (called makespan) of all jobs. With a modest degree of effort it is easy to see that this is but a dual-like version of a bin-packing problem. Specifying some limit (e.g., lower bound) on the overall completion time, we then seek an $m$-bin pack of chips (jobs) to the bins (processors) where each bin has capacity no greater than this specified limit on processing time for the analogous processor. When processors are nonidentical, the bin-packing interpretation reflects bin-dependent chip weights. The cost of bins corresponds to, say, a fixed, processor utilization charge (e.g., setup cost) and assuming that once a processor completes its processing assignment, any remaining idle time is costly (e.g., opportunity cost) we have the unused capacity penalty in (GBPP1).

Following, we present a brief review of bin packing and related literature after which, a rather extensive treatment of a series of models pertaining to (GBPP1) is included. Here, a variety of formulations are examined from the perspective of their potential usefulness. The most promising of these is explicitly demonstrated and then evaluated by a set of computational experiments.

### 1.2 Review of Relevant Literature

Presently, much of the literature deals only with the traditional bin-packing problem. Further, there appears to be less than complete standardization in the open literature regarding the terms bin-packing problem, cutting-stock problem, loading problem, and multiple-knapsack problem. Occasionally, the terms are used interchangeably, and often they are not. In order to avoid any confusion, the terms bin-packing or cutting-stock will be taken by convention here to mean the problem where all objects must be packed or cut and the objective is to minimize the number of bins or pieces of stock used. The terms loading or multiple knapsack will be taken to refer to the related problem where not all the objects must be loaded or assigned and the objective is to maximize the value of the loaded items.

One of the earliest appearances of the classic bin-packing problem in the literature is in the paper of Eilon and Christofides [6]. While several formulations of the bin-packing problem as well as related problems are discussed in [6], the one actually solved is as follows:

(6) Minimize: $$\sum_{i=1}^{m}\sum_{j=1}^{m} \mathrm{v}_j x_{ij}$$

(7) Subject to: $$\sum_{i=1}^{n} t_i x_{ij} \leqslant w, 1 \leqslant j \leqslant m$$

(8) $$\sum_{j=1}^{m} x_{ij} = 1, 1 \leqslant i \leqslant n$$

(9) $x_{ij}$ binary

where $\mathrm{v}_{j+1} = p\mathrm{v}_j$, $\mathrm{v}_1 = 1$, and $p$ is the largest number of objects that can be packed into a bin. This formulation is interesting because it eliminates the need for $y_j$ variables to indicate the use of a bin (each with capacity $w$) due to the structuring of the $\mathrm{v}_j$ coefficients which ensures that the smallest number of bins is used. Devised in [6] is an exact and a heuristic algorithm for solving the problem. The exact procedure is a variation of the tree-search algorithm of Balas [2] while the heuristic algorithm is a variation of the best-fit decreasing algorithm which will be discussed shortly.

Performance guarantees on the quality of the solution for bin-packing heuristics have received substantial attention. In 1972 and 1973 Garey, Graham and Ullman [10], presented worst-case results of various heuristic algorithms. However, it was Johnson's doctoral thesis in 1973 [19] which laid the theoretical groundwork for the worst-case analysis of the class of heuristic algorithms known as any-fit. A summary of this work appears in [20]. Also, Johnson, Demers, Ullman, Garey and Graham, have presented a comprehensive survey of both old and new results in [21].

The four any-fit algorithms of primary interest are as follows:

Let $0 = \{1, 2, \ldots, n\}$ be the set of objects or chips, and let
$J = \{1, 2, \ldots, m\}$ be the set of bins.

1. First-Fit (FF): Assign each object from set 0 in the order in which it appears in the set to the bin in set $J$ with the lowest index into which it will feasibly fit.

2. Best-Fit (BF): Assign each object from set 0 in the order in which it appears in the set to the bin in set $J$ such that the resulting unused capacity is minimal.

3. First-Fit Decreasing (FFD): Order the set 0 in nonincreasing order on the basis of object displacement. Apply the FF algorithm.

4. Best-Fit Decreasing (BFD): Order the set 0 in nonincreasing order on the basis of object displacement. Apply the BF algorithm.

The worst-case results are expressed as a ratio of the solution value of the heuristic to the optimal solution for each algorithm. The ratios for these algorithms, say $\bar{R}$, are given as asymptotic results and appear as follows [21]:

1. $\lim_{k\to\infty} \bar{R}_{FF}(k) = 17/10$

2. $\lim_{k\to\infty} \bar{R}_{BF}(k) = 17/10$

3. $\lim_{k\to\infty} \bar{R}_{FFD}(k) = 11/9$

4. $\lim_{k\to\infty} \bar{R}_{BFD}(k) = 11/9$

Here, $k$ is the optimum number of bins required for a given list of chip displacements.

More recently Hung and Brown in 1978 [16] considered a form of the bin-packing problem in which the bins were permitted to have differing capacities. Their formulation follows:

(10) Minimize: $\sum_{j=1}^{m} y_j$.

(11) Subject to: $\sum_{i=1}^{n} t_i x_{ij} \leqslant w_j y_j \qquad 1 \leqslant j \leqslant m$

(12) $\sum_{j=1}^{m} x_{ij} = 1 \qquad 1 \leqslant i \leqslant n$

(13) $x_{ij}$, $y_j$ binary.

The algorithm developed is basically one of implicit enumeration. However, a characterization of assignments was given which capitalizes on any equivalence in object displacements or bin capacities. This permitted a development of a set of rules which excludes redundant assignments.

In 1975 Ingargiola and Korsh [18] presented an enumerative algorithm for solving the related loading or multiple-knapsack problem. The formulation of the problem considered appears as:

$$\text{(14)} \quad \text{Minimize:} \quad \sum_{i=1}^{n} \sum_{j=1}^{m} c_i x_{ij}.$$

$$\text{(15)} \quad \text{Subject to:} \quad \sum_{i=1}^{n} t_i x_{ij} \leqslant w_j \qquad 1 \leqslant j \leqslant m$$

$$\text{(16)} \quad \sum_{j=1}^{m} x_{ij} \leqslant 1 \qquad 1 \leqslant i \leqslant n$$

$$\text{(17)} \quad x_{ij}, \text{ binary.}$$

The algorithm in [18] is based on the notion that before a search is attempted, as many decisions as possible should be made about the inclusion and exclusion of objects from the knapsacks. This is accomplished by the introduction of an ordering relation among objects which generates those objects that will be excluded in an optimal solution if a given object is included.

In 1978, Hung and Fisk [17] presented a branch-and-bound algorithm for solving the same problem as that of Ingargiola and Korsh. Here, it was shown that it is possible to compute the optimal LP dual multipliers in closed form. Furthermore, these multipliers for the first set of constraints were shown to equal a constant. Two relaxations of the problem were developed; lagrangian and surrogate. In the lagrangian relaxation the second set of constraints are relaxed by use of the optimal LP dual multipliers, and the problem decomposes into separate knapsack problems for each bin. In the surrogate relaxation the first set of constraints is combined into a single constraint, and the second set of constraints is substituted into this constraint; the problem then reduces to a single-knapsack problem.

In a follow-up article Fisk and Hung [9] presented a heuristic algorithm for solving the same problem. In fact, they used the same surrogate relaxation as discussed above. Their procedure is founded upon the observation that if a feasible solution can be found among the objects and bins identified in the optimal solution to the relaxed problem, then an optimal solution to the original problem has been found. The search for a feasible solution employs a set of exchange rules which considers all possible one-for-one, one-for-two and two-for-one exchanges of objects between bins.

Of additional interest is work relating to what has been popularized as "generalized assignment models." Included are a myriad of formulations which deviate in various ways from the traditional assignment theme. Subsumed by these generalizations, of course, is the classic bin-packing problem. A lucid exposition of an algorithmic developments within this area can be found in [24] and [25].

Other references pertaining to bin-packing and related problems can also be cited. A partial list would include works in [4, 5, 7, 11, 27].

## 2. METHODOLOGICAL DEVELOPMENTS

There are a variety of ways to approach (GBPP1), classic among which is a branch-and-bound scheme where bounds are derived from an appropriately constructed linear programming relaxation. In fact, such an approach was pursued by the authors where (GBPP1) is augmented with clique constraints

which tighten the relaxation. These clique constraints, derived from implicit as well as explicit packing restrictions, can be quite powerful in characterizing the integer hull (e.g., [23]) of (GBPP1) but can themselves be relatively difficult to construct. This should be obvious since determining cliques in graphs is NP-complete [12]. To this extent, the linear programming relaxations which were developed did not result in especially encouraging computational performance (especially when $R_j \neq \emptyset$). As a consequence, we turn to an alternative attack, also rather classic, where we create lagrangian relaxations of the problem [26]. In what follows, we consider various of these relaxations presenting computational results for the more promising ones.

Suppose we recast (GBPP1) in the form given by (GBPP2) below where we let $u_j = w_j y_j - \sum_{i=1}^{n} t_{ij} x_{ij}$, $z_j = 1 - y_j$ and where, in addition, we replace $u_j$ in the first set of constraints by a slack variable. Restating the problem as a maximization problem, we have:

$$\text{(18)} \quad \text{(GBPP2):} \quad \sum_{j=1}^{m} (c_j + d_j w_j) - \text{Maximize:} \quad \sum_{j=1}^{m} (c_j + d_j w_j) z_j + \sum_{j=1}^{m} \sum_{i=1}^{n} d_j t_{ij} x_{ij}.$$

$$\text{(19)} \quad \text{Subject to:} \quad \sum_{i=1}^{n} t_{ij} x_{ij} + w_j z_j \leqslant w_j, \quad 1 \leqslant j \leqslant m$$

$$\text{(20)} \quad \sum_{j=1}^{m} x_{ij} = 1, \qquad 1 \leqslant i \leqslant n$$

$$\text{(21)} \quad x_{ij},\ z_j \text{ binary.}$$

In the ensuing developments, we will adapt the notational convention whereby $q_j$ denotes the lagrange multiplier associated with the packing constraints and $r_i$, the multiplier associated with the assignment restrictions.

### 2.1 Lagrangian Relaxation of Packing Constraints

If the packing restrictions of (GBPP2) are relaxed, the lagrangian formulation for a given nonnegative vector $q$ can be stated by (LR1) below:

$$\text{(22)} \quad \text{(LR1):} \quad \sum_{j=1}^{m} (c_j + d_j w_j) - \sum_{j=1}^{m} w_j q_j - \text{Maximum:} \quad \sum_{j=1}^{n} (c_j + d_j w_j - w_j q_j) z_j + \sum_{j=1}^{m} \sum_{i=1}^{n} (d_j t_{ij} - t_{ij} q_j) x_{ij}.$$

$$\text{(23)} \quad \text{Subject to:} \quad \sum_{j=1}^{m} x_{ij} = 1, \quad 1 \leqslant i \leqslant n$$

$$\text{(24)} \quad x_{ij},\ z_j \text{ binary.}$$

Now, relaxing the packing constraints of (GBPP2) leads to a relaxation for which the optimal solution yields a bound value identical to that obtained from the solution of the LP relaxation of (GBPP2). This result follows directly from the integrality property given in Geoffrion [13] which states that an optimal value of a lagrangian relaxation is not altered by dropping the integrality conditions on its variables. The implication of this property is fairly immediate and can be given by the following:

THEOREM 1: Let the LP relaxation be feasible and let the lagrangian relaxation have the integrality property. Then the maximum value of the lagrangian relaxation is equal to the value of the LP relaxation.

PROOF: The theorem follows trivially from the integrality property specified in [13].

Clearly, (LR1) possesses the integrality property and by Theorem 1 its maximum value equals the value of the LP relaxation of (GBPP2). As a consequence, little may be gained in formulating and solving (LR1) that is not otherwise achievable by simply solving the LP relaxation of (GBPP2). For this reason, the model relaxing the packing constraints is not pursued further.

### 2.2 Lagrangian Relaxation of the Assignment Constraints

Suppose we pursue an alternative lagrangian formulation based upon a relaxation of the constraints pertaining to assignment. As it turns out, this relaxation will prove to be the most fruitful in this work and, in fact, provides the basis of the principle algorithm developed and subsequently tested.

If the assignment constraints are relaxed, the lagrangian relaxation of (GBPP2) for a given vector $r$ can be given by (LR2) below:

$$(25)\quad \text{(LR2)}: \sum_{j=1}^{m}(c_j + d_j w_j) - \sum_{i}^{n} r_i - \quad \text{Maximum:} \quad \sum_{j=1}^{m}(c_j + d_j w_j) z_j + \sum_{j=1}^{m}\sum_{i=1}^{n}(d_j t_{ij} - r_i) x_{ij}$$

$$(26)\quad \text{Subject to:} \quad \sum_{i=1}^{n} t_{ij} x_{ij} + w_j z_j \leqslant w_j,\ 1 \leqslant j \leqslant m$$

$$(27)\quad x_{ij},\ z_j \text{ binary.}$$

To maximize the value of this relaxation, we want to find those multipliers $r_i$ which solve the following lagrangian dual problem:

$$(28)\quad \text{(LD1)}: \text{Minimize:} \quad \sum_{i=1}^{n} r_i + \left\{\text{maximum:} \sum_{j=1}^{m}(c_j + d_j w_j) z_j + \sum_{j=1}^{m}\sum_{i=1}^{n}(d_j t_{ij} - r_j) x_{ij} \;\middle|\; \sum_{i=1}^{n} t_{ij} x_{ij} + w_j z_j \leqslant w_j,\ 1 \leqslant j \leqslant m,\ x_{ij},\ z_j \text{ binary}\right\}$$

$$(29)\quad \text{Subject to: } r_i \text{ unrestricted, } 1 \leqslant i \leqslant n.$$

Consider the maximization part of (LD1). The following result defines an optimality condition that must hold for this maximization problem:

LEMMA 1: Let $z^*$, $x^*$ be the optimal solution to (LD1) for a given $r$. Then

$$(30)\quad (c_j + d_j w_j) z_j^* + \sum_{i}(d_j t_{ij} - r_i) x_{ij}^* \geqslant (c_j + d_j w_j), \quad 1 \leqslant j \leqslant m.$$

PROOF: By contradiction suppose that at optimality (30) is violated for some bin $j$. This implies that $z_j^* = 0$ and that $x_{ij}^* = 1$ for one or more objects. However, the value of this solution can always be increased by letting $z_j^* = 1$ $x_{ij}^* = 0$ for all $i$ which contradicts the assumption of optimality and the proof is complete.

By use of Lemma 1, (LD1) can be transformed into the following problem:

(31) (LD2): Minimize: $\sum_{i=1}^{n} r_i + \sum_{j=1}^{m} q_j + \sum_{j=1}^{m} (c_j + d_j w_j)$

(32) Subject to: $q_j \geqslant \sum_{i \in I_{jk}} (d_j t_{ij} - r_i) - (c_j + d_j w_j), \quad 1 \leqslant k \leqslant p_j$

$1 \leqslant j \leqslant m$

(33) $q_j \geqslant 0,\ 1 \leqslant j \leqslant m;\ r_i$ unrestricted, $\quad 1 \leqslant i \leqslant n.$

where $I_{jk}$ is the set of objects in the $k$th feasible pack for bin $j$ and $p_j$ is the number of feasible packs for bin $j$.

To see the relationship creating (LD2) note that the objective function of the maximization part of (LD1) must contain a constant term, $c_j + d_j w_j$ for each bin $j$. The value of the objective function in excess of this constant is $\sum_{i \in I_{jk}} (d_j t_{ij} - r_i) - (c_j + d_j w_j)$ since the excess can only occur when objects are packed into bin $j$ and not when bin $j$ is closed. Furthermore, the objective function of (LD1) requires that the value in excess of this constant be maximized. If we let $q_j$ represent the maximum value then $q_j \geqslant \sum_{i \in I_{jk}} (d_j t_{ij} - r_i) - (c_j + d_j w_j) \geqslant 0$ for all $k$ which is the constraint set and the sign restriction for $q_j$ in (LD2). Finally, since the objective function of (LD1) requires that $\sum_{j=1}^{m} q_j$ be minimized, we have the objective function of (LD2).

Note that the objective function of (LD2) contains a constant term which cancels exactly the constant term in the objective function of (LR2). Since (LD2) is a transformation of (LD1), the lagrangian dual problem for (LR2) can be written as follows:

(34) (LD3): Minimize: $\sum_{i=1}^{n} r_i + \sum_{j=1}^{m} q_j.$

(35) Subject to: $\sum_{i \in I_{jk}} r_i + q_j \geqslant \sum_{i \in I_{jk}} d_j t_{ij} - (c_j + d_j w_j), \quad 1 \leqslant k \leqslant p_j$

$1 \leqslant j \leqslant m$

(36) $q_j \geqslant 0,\ 1 \leqslant j \leqslant m;\ r_i$ unrestricted, $\quad 1 \leqslant i \leqslant n\ .$

Writing the LP dual of this problem we have (LD4):

(37) (LD4): Minimize: $\sum_{jk} \left[ c_j + d_j w_j - \sum_{i \in I_{jk}} d_j t_{ij} \right] v_{jk}\ .$

(38) Subject to: $\sum_{jk \in I_i} v_{jk} = 1,\ \ 1 \leqslant i \leqslant n$

(39) $\sum_{k} v_{jk} \leqslant 1\ ,\ \ 1 \leqslant j \leqslant m$

(40) $$v_{jk} \geqslant 0, \qquad 1 \leqslant k \leqslant p_j, \quad 1 \leqslant j \leqslant m.$$

where $v_{jk}$ is the dual multiplier associated with each constraint in (LD3), the objective function has been changed from a maximization to a minimization, and $I_i = \{(j, k) | i \in I_{jk}\}$.

Observe that due to the constraints, each $v_{jk}$ in (LD4) is further restricted to be upper-bounded by 1. If each $v_{jk}$ is thought of as a variable which selects a pack $I_{jk}$ or some portion of it, then (LD4) is the problem of selecting a set of packs (possibly fractional) such that the cost of the packs is minimized, each object is packed once, and at most one pack is selected from each bin. The cost of a pack is the cost of a bin, $c_j$ plus the cost of any unused capacity in bin $j$, $d_j w_j - \sum_{i \in I_{jk}} d_j t_{ij}$.

Problem (LD4) can be further constrained by the addition of a lower bound on the number of packs that must be selected. With the addition of this constraint, (LD4) can be rewritten as follows:

(41) (LD5): Minimize: $$\sum_{jk} \left( c_j + d_j w_j - \sum_{i \in I_{jk}} d_j t_{ij} \right) v_{jk}.$$

(42) Subject to: $$\sum_{jk \in I_i} v_{jk} = 1, \quad 1 \leqslant i \leqslant n$$

(43) $$\sum_{k} v_{jk} \leqslant 1, \quad 1 \leqslant j \leqslant m$$

(44) $$\sum_{jk} v_{jk} \geqslant l \text{ and } v_{jk} \geqslant 0, \qquad 1 \leqslant k \leqslant p_j, \quad 1 \leqslant j \leqslant m.$$

Here, $l$ is a lower bound on the number of bins required.

Relaxing the assignment constraints leads to a lagrangian relaxation for which determining the optimal value of the relaxation is equivalent to solving a large LP problem. This notion, with the exception of the constraint on the minimum number of bins, follows from the results of Fisher, Northup, and Shapiro [8]. The discrete optimization problem which is considered takes the following form:

(45) Minimize: $f(\underline{x})$

Subject to: $g(\underline{x}) \leqslant \underline{b}$

$\underline{x} \in X$

where $f(\underline{x})$ is a scalar valued function, $g(\underline{x})$ is a function from $R^n$ to $R^m$, and $X$ is a set of discrete elements. Fisher showed that the lagrangian dual of this problem, where the constraints are relaxed, can be transformed into the following problem:

(46) Minimize: $$\sum_{t=1}^{T} \lambda_t f(\underline{x}^t).$$

Subject to: $$\sum_{t=1}^{T} \lambda_t g(\underline{x}^t) \leqslant b$$

$$\lambda_t \geqslant 0 \text{ for all } t$$

where $\underline{x}^t$ is an element in the set $X$ and $T = |X|$.

The equivalence between (LD4) and the above LP is easy to verify:

1. The index $k$ is equivalent to $t$. When the assignment constraints are relaxed, (GBPP2) decomposes into separate problems for each bin. Thus, there is a separate index $k$ for each bin $j$.

2. $X_j = \{I_{j1}, I_{j2}, \ldots, I_{jk}\}$, the set of feasible packs for bin $j$, is equivalent to $X$.

3. $I_{jk}$ is equivalent to $\underline{x}^t$. That is, $I_{jk}$ is an element of $X_j$.

4. $\left(c_j + d_j w_j - \sum_{i \in I_{jk}} d_j t_{ij}\right)$ is equivalent to $f(\underline{x}^t)$.

5. $\sum_{i \in I_{jk}} v_{jk}$ is equivalent to $g(\underline{x}^t)$ and $\underline{1}$ is equivalent to $\underline{b}$.

6. The convexification constraint in (LD4) can also be stated as an equality if one includes the empty pack where $y_j = 0$ as a feasible pack. The variable representing the empty pack is the slack variable.

7. $v_{jk}$ is equivalent to $\lambda_t$.

Let us return to (LD5). If one were to enumerate all feasible packs for every bin and solve the resulting LP problem, the amount of computational effort would be large indeed. However, it may not be necessary to consider all possible packs or for that matter even a large number of them. What are of interest are the "goods" packs; that is, those which will be in an optimal solution to (LD5). Required then is a procedure for generating these "goods" packs. Such a procedure can be stated as follows:

*A Column Generation Procedure*

Step (0): Generate a slack variable for each bin and add it to the problem. Determine a value for $l$, the minimum number of bins required. If a feasible solution to (GBPP2) is known, generate the packs which correspond to this solution and add these columns to the problem.

Step (1): Solve (LD5) yielding a set of optimal dual multipliers. Let $r^*$ correspond to the first set of constraints in (LD5), $q^*$, the second set, and $s^*$, the third set.

Step (2): For each bin $j$, solve the following knapsack problem:

(47) Maximize: $\sum_{j=1}^{m} (d_j t_{ij} + r_i^*) x_{ij}$

(48) Subject to: $\sum_{j=1}^{m} t_{ij} x_{ij} \leqslant w_j$.

Let $x_j^*$ be the optimal solution for bin $j$.

Step (3): For each bin $j$, check if the column generated is an improving column. That is, if

(49) $$\sum_i r_i^* x_{ij}^* + q_j^* + s^* > c_j + d_j w_j - \sum_i d_j t_{ij} x_{ij}^*$$

then add the column corresponding to $x_j^*$ to the problem.

Step (4): If at least one of the generated columns is an improving column, go to Step 1. Otherwise, the current solution to (LD5) is optimal. If the current solution to (LD5) contains artificial variables, then there exists no feasible solution to (GBPP2).

The framework of the procedure above is, of course, standard and has enjoyed frequent use elsewhere. In particular, the use of a column generation scheme was employed in [14, 15] for the solution of the classic cutting-stock problem.

### 2.3 The Development of the Branch-and-Bound Procedure

In this section, we suggest a branch-and-bound procedure which is used in resolving any differences between the solution to (LD5) and the solution to (GBPP2). Since the LP problem in (LD5) results from a langrangian relaxation of the assignment constraints, it is possible that the value of the optimal solution to (LD5) may be less than the value of the optimal solution to (GBPP2). Furthermore, even if the value of the optimal solution to (LD5) equals that of the optimal solution to (GBPP2), the solution to (LD5) may be fractional.

Bounds in the branch-and-bound procedure are determined by solving a relaxed problem with the optimal dual multipliers from the solution to (LD5). The problem which is relaxed is as follows:

(50) (GBPP3): Minimize: $\sum_{j=1}^{m} (c_j + d_j w_j) y_j - \sum_{j=1}^{m} \sum_{i=1}^{n} d_j t_{ij} x_{ij}$

(51) Subject to: $\sum_{i=1}^{n} t_{ij} x_{ij} \leqslant w_j y_j, \qquad 1 \leqslant j \leqslant m$

(52) $\sum_{j=1}^{m} x_{ij} = 1 \qquad , \qquad 1 \leqslant i \leqslant n$

(53) $x_{ij}, \ y_j$ binary.

Problem (GBPP3) follows from (GBPP1) by letting $u_j = w_j y_j - \sum_{i=1}^{u} t_{ij} x_{ij}$ in the objective function and by replacing $u_j$ with a slack variable in the first set of constraints. The lagrangian relaxation of the assignment constraints of (GBPP3) for a given $r$ result in (LR3) below:

(54) (LR3): $\sum_{i=1}^{n} r_i +$ Minimize: $\sum_{j=1}^{m} (c_j + d_j w_j) y_j - \sum_{j=1}^{m} \sum_{i=1}^{n} (d_j t_{ij} + r_i) x_{ij}$

(55) Subject to: $\sum_{i=1}^{n} t_{ij} x_{ij} \leqslant w_j y_j, \quad 1 \leqslant j \leqslant m$

(56) $x_{ij}, \ y_j$ binary.

In this form, (LR3) contains no constraint on the number of bins that are required to solve the problem. The relaxation could be augmented, for example, with a constraint of the form $\sum_{j=1}^{m} y_j \geqslant k$ where $k$ is a lower bound on the number of bins required. However, this constraint makes no distinction between the capacity of the bins. A constraint of the following form $\sum_{j=1}^{m} w_j y_j \geqslant t$ may be more discriminating. Here $t$ is a lower bound on the bin capacity needed to solve the problem. If the displacement of objects were independent of the bins in which the objects are packed, then $t$ would simply

be $\sum_{i=1}^{n} t_i$. However, the displacement of objects can vary from bin to bin; thus, $t$ must be determined as $\sum_{i=1}^{n} \min_{1 \leq j \leq m} (t_{ij})$, since there may exist a solution where each object attains its minimum displacement. However, depending on the range of object displacements, this capacity constraint may be rather weak. A capacity constraint which seems to more accurately reflect the capacities of the bins can be given as $\sum_{j=1}^{m} \bar{a}_j y_j \geqslant n$ where $\bar{a}_j$ is the maximum number of objects that can be packed into bin $j$. Clearly this constraint is valid since any feasible solution must allow for $n$ objects to be packed.

The value of $\bar{a}_j$ for each bin $j$ can be determined by solving the following knapsack problem (KP1):

(57) (KP1): Maximize: $\sum_{i=1}^{n} x_{ij} = \bar{a}_j$

(58) Subject to: $\sum_{i=1}^{n} t_{ij} x_{ij} \leqslant w_j$

(59) $x_{ij}$ binary.

While this problem can be treated with any standard knapsack algorithm, it can also be solved by the following procedure:

*KP1 Solution*

Step (1): For bin $j$ sort the objects in nondecreasing order on the basis of $t_{ij}$. Let (i) denote the $i$ th element in the sort.

Step (2): Let $p$ be the greatest integer $(0 \leqslant p \leqslant n)$ such that $\sum_{i=1}^{p} t_{(i)j} \leqslant w_j$.

Step (3): The optimal solution is $x_{(i)j} = 1$ for $i = 1, 2, \ldots, p$ and $x_{(i)j} = 0$ for $i = p + 1, p + 2, \ldots, n$ if $p > 0$ or $x_{ij} = 0$ for all $i$ if $p = 0$.

LEMMA 2: The KP1 solution procedure solves problem (KP1) optimally.

PROOF: Let $p^*$ be the value of $p$ determined in Step 2 of the procedure. If $p^* = 0$, then the solution is trivially optimal. If $p^* > 0$ then it is possible to pack at least $p^*$ chips into bin $j$ as shown in Step 3. Let the chips packed be denoted by $(1), (2), \ldots, (p^*)$. The chips not packed are then $(p^* + 1), (p^* + 2), \ldots, (n)$ such that $t_{(i)j} > u_j$ for $i = p^* + 1, p^* + 2, \ldots, n$ where $u_j = w_j - \sum_{i=1}^{p^*} t_{(i)j}$. In order for a solution to (KP1) to exist in which it is possible to pack more than $p^*$ chips into bin $j$, it must be possible to exchange at least $q + 1$ chips not in the current pack for $q$ chips which are in the pack. However, the total displacement of any $q$ chips not packed must be at least equal to the displacement of any $q$ which are packed. Further, the displacement of any size chip not packed must exceed $u_j$. Thus, it is not possible to create an exchange of $q + 1$ unpacked chips for $q$ which are packed into bin $j$. Hence, $p^*$ is the maximum number of chips which can be packed and the proof is complete.

When augmented with the capacity constraint, (LR3) can be rewritten in the form given by (LR4):

(60) (LR4): $\sum_{i=1}^{n} r_i +$ Minimize: $\sum_{j=1}^{m} (c_j + d_j w_j) - \sum_{j=1}^{m} \sum_{i=1}^{n} (d_j t_{ij} + r_i) x_{ij}$

(61) Subject to: $\sum_{i=1}^{n} t_{ij} x_{ij} \leqslant w_j y_j \quad 1 \leqslant j \leqslant m$

(62) $\sum_{j=1}^{m} \bar{a}_j y_j \geqslant n$

(63) $x_{ij}$, $y_j$ binary.

The solution to (LR4) can be obtained as follows:

*LR4 Solution Procedure*

Step (1): For a given $r$, solve the following knapsack problem for each bin:

(64) Maximize: $\sum_{j=1}^{m} (d_j t_{ij} + r_i) x_{ij}$

(65) Subject to: $\sum_{i=1}^{n} t_{ij} x_{ij} \leqslant w_j$

(66) $x_{ij}$ binary.

Let $x^*$ be the optimal solution to this problem for bin $j$ and let $v_j = c_j + d_j w_j - \sum_{i=1}^{n} (d_j t_{ij} + r_i) x_{ij}^*$.

Step (2): For each bin $j$ solve KP1. Let $x_j'$ be the optimal solution to (KP1) and let $\bar{a}_j = \sum_i x_{ij}'$.

Step (3): Solve the following knapsack problem:

(67) Minimize: $\sum_{j=1}^{m} v_j y_j$

(68) Subject to: $\sum_{j=1}^{m} \bar{a}_j y_j \geqslant n$

(69) $y_j \in \{0, 1\}$.

Let $y^*$ be the optimal solution to this problem and let the value of the lagrangean relaxation $L = \sum_{i=1}^{n} r_i + \sum_{j=1}^{m} v_j y_j^*$.

The correctness of this procedure is given by the following result:

THEOREM 3: The LR4 solution procedure solves problem (LR4).

PROOF: In an optimal solution to (LR4) either $y_j = 0$ or $y_j = 1$ for each $j$. If $y_j = 0$ for some $j$, then $x_{ij} = 0$ for all $i$ and the value of the lagrangian relaxation associated with bin $j$ is zero (i.e., $v_j = 0$). However, if $y_j = 1$, then the optimal value of the lagrangian relaxation associated with bin $j$ is the optimal value of the following problem:

$$\text{(70)} \qquad \text{Minimize:} \quad (c_j + d_j w_j) - \sum_{i=1}^{n} (d_j t_{ij} + r_i) x_{ij}$$

$$\text{(71)} \qquad \text{Subject to:} \quad \sum_{i=1}^{n} t_{ij} x_{ij} \leqslant w_j$$

$$\text{(72)} \qquad x_{ij} \in \{0, 1\}.$$

This, however, is the same problem as that which is solved in Step (1) of the (LR4) solution procedure. Thus, $v_j$ is the optimal value of the lagrangian relaxation asociated with bin $j$ and the solution of (LR4) reduces to the solution of the same problem which is solved in Step (3) of the procedure.

The branch-and-bound procedure to be used in resolving differences between the optimal solution to (LD5) and the optimal solution to (GBPP2) can now be stated formally.

*Solution Procedure for (GBPP2)*

Step (0): Solve (LD5): if the optimal solution is integer, stop. Let the subproblem list be empty, and let $p = 0$. Let the candidate bound $C^*$ be a large positive value. Perform Steps 1 through 4 and go to Step 5.

Step (1): Using the optimal lagrangian multipliers from Step 0. Solve (LR4) to yield a current bound $C$ and a current solution $(x, y)$.
If $C \geqslant C^*$ go to Step 4.

Step (2): If for the current solution

$$\sum_{j \in J} x_{ij} \neq 1 \text{ for any } i$$

where $J = \{j | y_j = 1\}$

go to Step 3.

Update the candidate bound and solution $(x^*, y^*)$ with the current bound and solution and remove all subproblems from the subproblem list with a bound equal to or greater than the candidate bound. Go to Step 4.

Step (3): Let $I_1 = \left\{ i \,\middle|\, \sum_{j \in J} x_{ij} \geqslant 1 \right\}$. $I_2 = \left\{ i \,\middle|\, \sum_{j \in J} x_{ij} = 0 \right\}$

If $I_1 = \emptyset$ go to Step 3.1
Let $J_i = \{j | j \in J,\ i \in I_1, \text{ and } x_{ij} = 1\}$
Select for branching the variable $x_{kl}$ such that

$$\theta_{kl} = \max_{i \in I_1} \{\max_{j \in J_i} \{(d_j t_{ij} + r_i^*)/t_{ij}\}\}.$$

Add the subproblem to the subproblem list and go to Step 4.

3.1 Let $J_i' = \{j | j \in J$ and object $i \in I_2$ is not prohibited from bin $j\}$. If $J_i' = \emptyset$ for all $i \in I_2$ go to Step 3.2. Select for branching the variable $x_{kl}$ such that

$$\theta_{kl} = \max_{i \in I_2} \max_{j \in J_i'} \{(d_j t_{ij} + r_i^*)/t_{ij}\}$$

Add the subproblem to the subproblem list and to to Step 4.

3.2 Let $J_i'' = \{j | j \notin J$ and object $i \in I_2$ is not prohibited from bin $j\}$. If $J_i'' = \emptyset$ for any $i$, a feasible solution does not exist; go to Step 4. Select for branching the variable $x_{kl}$ such that

$$\theta_{kl} = \min_{i \in I_2} \min_{j \in J_i''} \{v_j/\bar{a}_j\}.$$

Add the subproblem to the subproblem list and to to Step 4.

Step (4): Set $p \leftarrow p + 1$.

Step (5): Select from the subproblem list the subproblem with the least lower bound. If the subproblem list is empty, stop.

Step (6): Let $x_{kl} = 0$ and perform Step 1 through 4. Let $x_{kl} = 1$ and $y_l = 1$ and perform Step 1 through 4. Go to Step 5.

### 2.4 An Approximate Langrangian Relaxation

Recall, from the branch-and-bound procedure for (GBPP2) just presented, that the optimal multipliers are first determined by column generation procedure. The branch-and-bound scheme is then invoked to determine the optimal primal solution. We could, of course, have employed a subgradient approach to determine these multipliers. Indeed such an attack was examined; however, the usual computational difficulties arose (e.g., step size adjustment, etc.) and the procedure was abandoned after a preliminary investigation. Regardless, with either the column generation or the subgradient procedure no further adjustment of the lagrange multipliers is pursued after the initial optimization due to, in fact, the computational expense incurred. Suppose, however, it were possible to approximate the values of the optimal multipliers in a manner that was not computationally expensive. In such a case, it might be possible to recompute multipliers at each node in the branch-and-bound tree without sustaining prohibitive computational effort. Following, we develop precisely such an approach.

Clearly, one method for approximating the multiplier values would be to simply use the optimal values of the dual multipliers from the LP relaxation. It is the case however, that in those instances where the cost of unused capacity dominates or affects significantly the optimal solution value, the bounds derived by using the optimal LP dual multipliers for an approximation can be quite weak. We must have another alternative and as such, consider again (LR4). In order to maximize the value of this relaxation, we seek those multipliers $r$ which solve the lagrangian dual, (LD6) below:

$$(73) \quad \text{(LD6): Maximize:} \quad \sum_i r_i + \left\{ \text{Minimum} \sum_{j=1}^{m} (c_j + d_j w_j) y_j - \sum_{j=1}^{m} \sum_{i=1}^{n} (d_j t_{ij} + r_i) x_{ij} \right|$$

$$\text{Subject to: } r_i \text{ unrestricted}_{1 \leqslant i \leqslant n} \qquad \left.\begin{array}{l} \displaystyle\sum_{i=1}^{n} t_{ij}x_{ij} \leqslant w_j y_j, \quad 1 \leqslant j \leqslant m \\ \\ \displaystyle\sum_{j=1}^{n} \bar{a}_j y_j \geqslant n \\ \\ x_{ij},\ y_j \text{ binary} \end{array}\right\}.$$

It is well known that the objective function of this problem is concave in the $r_i$ variables [9].

Suppose now that all the lagrange multipliers are constrained to be equal in value. Then the problem can be restated as follows:

$$\text{(74)} \quad \text{(LD7): Maximize:} \quad \sum_{i=1}^{n} r_i + \left\{ \text{Minimum} \sum_{j=1}^{m} (c_j + c_j w_j) y_j - \sum_{j=1}^{m} \sum_{i=1}^{n} (d_j t_{ij} + r_i) x_{ij} \right|$$

$$\text{Subject to: } r_i = r \text{ and unrestricted}_{\text{for } 1 \leqslant i \leqslant n} \qquad \left.\begin{array}{l} \displaystyle\sum_{i=1}^{n} t_{ij}x_{ij} \leqslant w_j y_j, \quad 1 \leqslant j \leqslant m \\ \\ \displaystyle\sum_{j=1}^{m} \bar{a}_j y_j \geqslant n \\ \\ x_{ij},\ y_j \text{ binary} \end{array}\right\}.$$

or equivalently,

$$\text{(75)} \quad \text{Maximize:} \quad nr + \left\{ \text{Minimum} \sum_{j=1}^{m} (c_j + d_j w_j) y_j - \sum_{j=1}^{m} \sum_{i=1}^{n} (d_j t_{ij} + r_i) x_{ij} \right|$$

$$\text{Subject to: } r \text{ unrestricted} \qquad \left.\begin{array}{l} \displaystyle\sum_{i=1}^{n} t_{ij}x_{ij} \leqslant w_j y_j, \quad 1 \leqslant j \leqslant m \\ \\ \displaystyle\sum_{j=1}^{n} \bar{a}_j y_j \geqslant n \\ \\ x_{ij},\ y_j \text{ binary} \end{array}\right\}.$$

A constraint set for the $r_i$ variables has been introduced into the problem. However, these constraints are all convex and so, the lagrangian dual now pertains to the maximization of a concave function over a convex set. Let $L^*$ be the optimal value of the lagrangian relaxation without the constraints on $r_i$ and $\bar{L}^*$, the optimal value with the constraints on $r_i$. Since the latter problem is constrained and the former is not, then it must be that $\bar{L}^* \leqslant L^*$.

The restrictions on the values of the $r_i$ variables will only be worthwhile if (LD7) is relatively easy to solve. In order to show how to solve this problem, we first need to establish some properties for (LD7). The concavity of the objective function in $r$ has already been established. The following property of the objective function also holds:

PROPERTY 1: The objective function of (LD7) is piecewise linear in $r$.

PROOF: This property follows from the minimization portion of the objective function of (LD7). Let $r$ have the value $r'$. Then there exists an optimal solution to the minimization part of (LD7). Call this solution $(\underline{x}', \underline{y}')$ and let $r'$ change by an amount $\Delta$. For $\Delta$ small enough, the optimal solution to the minimization problem will remain unchanged. The change in $r'$ will cause the objective function of (LD7) to change by an amount $\left[n - \sum_{j=1}^{m}\sum_{i=1}^{n} x'_{ij}\right]\Delta$ which is linear in $\Delta$. If $\Delta$ is made large enough, then the optimal solution to the minimization problem will change and the objective function will be defined by a new line.

PROPERTY 2: The slope of the ojective function of (LD7) at any point $r'$ is

$$n - \sum_{j=1}^{m}\sum_{i=1}^{n} x'_{ij}$$

where $(x', y')$ is the optimal solution to the minimization problem in (LD7) at $r = r'$.

PROOF: This result follows directly from Property 1.

The plot of $\bar{L}$ versus $r$ for a typical problem might appear as shown in Figure 1. The line segments are numbered 1 through 7, and the optimal value $\bar{L}^*$ occurs at the intersection of line segments 3 and 4. Since the objective function of (LD7) is piecewise linear and concave in $r$, the optimal value will always occur at either the intersection of two line segments or anywhere along a line segment whose slope is zero. This suggests the following procedure for finding the value of $r$ which optimally solves (LD7). We note that other alternatives exist for the solution of (LD7).

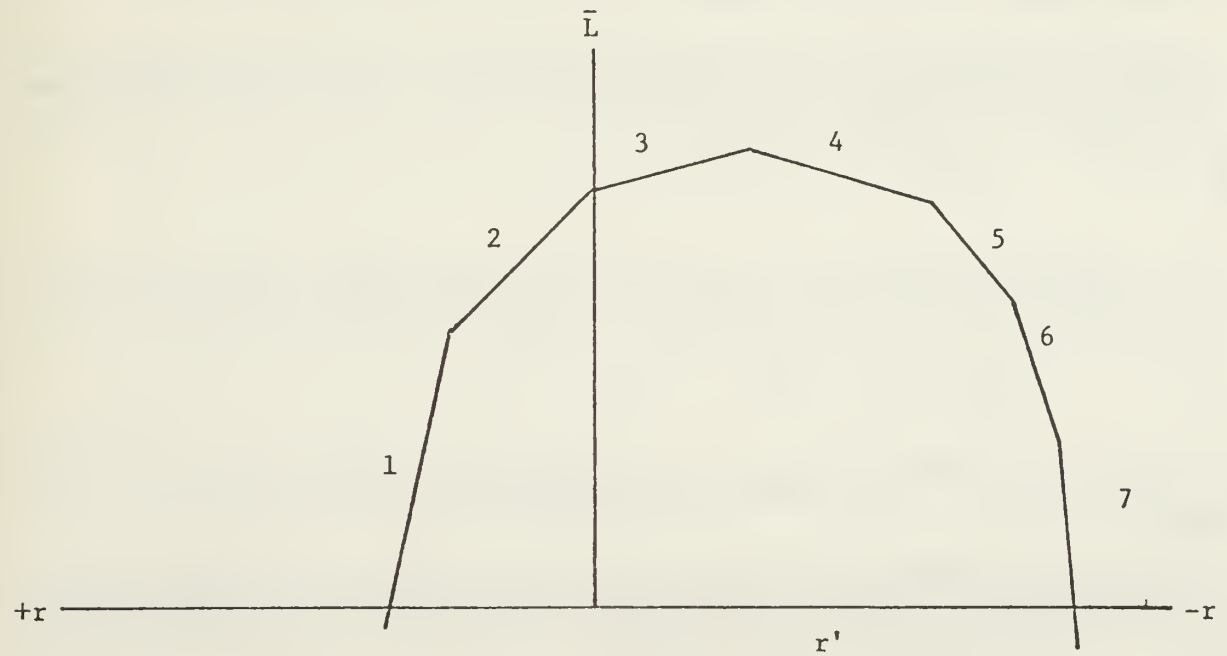



FIGURE 1. Solution value of (LD7) versus $r$.

*Solution of (LD7)*

Step (0): Select an initial value for $\epsilon > 0$. Let $m_- = m_+ = r = 0$.

Step (1): Solve (LR4) with $r_i = r$ for all $i$. Let $m$ be the slope of the objective function of (LD7) at $r$, and let $i$ be its $\bar{L}$ intercept. If $m = 0$ stop. If $m > 0$, go to Step 3.

Step (2): Let $m_- = m$, $i_- = i$, and $r_- = r$. If $m_+ \neq 0$, go to Step 4. Let $r = r - \lambda$ where $\lambda > 0$ is a step size. Go to Step 1.

Step (3): Let $m_+ = m$, $i_+ = i$, and $r_+ = r$. If $m_- \neq 0$, go to Step 4. Let $r = r + \lambda$ where $\lambda > 0$ is a step size. Go to Step 1.

Step (4): Let $r'$ be the point at which the two lines defined by $m_+$, $i_+$ and $m_-$, $i_-$ intersect. If $|r - r'| < \epsilon$ stop. Let $r = r'$, and go to Step 1.

We can give the following theorem regarding the solution determined by the algorithm:

THEOREM 4: The LD7 solution procedure terminates with a value of $r$ which is at most a distance $\epsilon > 0$ from $r^*$ which optimally solves (LD7).

PROOF: Clearly the procedure terminates. Either a line segment with a slope of zero is found; or since $\lambda > 0$, eventually a point $r_+$ is found on one side of $r^*$ and a point $r_-$ is found on the other side. The interval $(r_+ - r_-)$ is successively reduced in magnitude to value of $\epsilon$ or less. If the procedure terminates in Step 1, then a line segment for the objective function has been found which has a slope of zero at $r$. Any point on this line segment will yield an optimal solution to (LD7). If the procedure terminates in Step 4, then $r$ is either on the line defined by $m_+$, $i_+$ or $m_-$, $i_-$. Without loss of generality assume that $r$ lies on the line segment defined by $m_+$, $i_+$. Then at the point $r$ the objective function of (LD7) slopes downward in the direction of negative $r$, and at the point $r'$ which is at most a distance $\epsilon$ from $r$ the objective function of (LD7) slopes downward in the direction of positive $r$. Thus the value of $r$ at termination is at most a distance $\epsilon$ from $r^*$ which optimally solves (LD7).

Before this procedure can be implemented, it is first necessary to devise some means of determining the step size $\lambda$. If $r$ is negative, then the sign of the terms $(d_j t_{ij} + r)$ can be used to gauge the step size. For example, if the slope is negative, then there are too many objects assigned and the number must be reduced. Let $\lambda$ be determined as follows:

$$\lambda = \min_{i,j} \{(d_j t_{ij} + r) | (d_j t_{ij} + r) > 0\}. \tag{76}$$

Then $\lambda$ will be the minimum decrease in $r$ such that at least one $x_{ij}$ coefficient will no longer be positive. If one of these objects was previously assigned, then it would no longer be attractive to assign it and the number of assigned objects would decrease. Conversely, if the slope is positive, then there are too few objects assigned and the number must be increased. Let $\lambda$ be determined as follows:

$$\lambda = \min_{i,j} \{-(d_j t_{ij} + r) | (d_j t_{ij} + r) < 0\}. \tag{77}$$

Then $\lambda$ will be the minimum increase in $r$ such that at least one $x_{ij}$ coefficient will no longer be negative. Now at least one of these objects can be assigned and the number of assigned objects can increase.

If $r$ is positive, then the sign of the terms $(d_j t_{ij} + r)$ can no longer be used to gauge the step size since the sign will always be positive. Instead, if the slope is positive, let $\lambda$ be determined as follows:

$$\lambda = \min_{j} \{d_j/(\bar{a}_j - a_j) | (\bar{a}_j - a_j) \neq 0\} \tag{78}$$

where: $\bar{a}_j$ = the maximum number of objects that can be assigned to bin $j$
$a_j$ = the current number of objects that are assigned to bin $j$.

Since there are too few objects assigned and the number must be increased, $\lambda$ is the minimum increase in $r$ such that the number of objects assigned will equal the maximum number that can be assigned for some bin. To see this, let $v_j = i_j - a_j r$ and $\bar{v}_j = \bar{i}_j - \bar{a}_j r$ where $v_j$ is the value of the current solution for bin $j$, $i_j$ is its value when $r = 0$, $\bar{v}_j$ is the value of the solution for bin $j$ when the maximum number of objects are assigned to bin $j$, and $\bar{i}_j$ is its value when $r = 0$. These two solutions have the same value (i.e., $v_j = \bar{v}_j$) when $r = (\bar{i}_j - i_j)/(\bar{a}_j - a_j)$. If these two solutions are different (i.e., $\bar{a}_j \neq a_j$), then $\bar{i}_j$ and $i_j$ must differ by at least the cost of one unit of unused capacity (i.e., $\bar{i}_j = i_j + d_j$). Thus, $r$ must increase by at least $d_j/(\bar{a}_j - a_j)$.

If $r$ is positive and the slope is negative, let $\lambda$ be determined by setting $\lambda = r$. Since there are too many objects assigned, the number must be reduced. However, it does not seem possible to determine a minimum decrease in $r$ such that the number of assigned objects will decrease. Setting $r$ equal to zero, which is what this step size does, will certainly not increase the number of assigned objects and will probably decrease the number substantially.

With the addition of the step size calculations, the procedure for solving (LD7) can be restated as follows:

*Updated Solution Procedure for (LD7)*

Step (0): Select an initial value for $\epsilon > 0$. Let $r = m_+ = m_- = 0$.

Step (1): Solve LR4 with $r_i = r$ for all $i$. Let $m$ be the slope of the objective function of (LD7) at $r$, and let $i$ be its $\bar{L}$ intercept. If $m = 0$, stop. If $r > 0$, go to Step 3.

Step (2): If $m > 0$, go to Step 2.1. Let $m_- = m$ and $i_- = i$. If $m_+ \neq 0$, go to Step 2.2. Let $\lambda = \min_{i,j} \{(d_j t_{ij} + r): (d_j t_{ij} + r) > 0\}$. Let $r = r - \lambda$, and go to Step 1.

2.1 Let $m_+ = m$ and $i_+ = i$. If $m_- \neq 0$, go to Step 2.2. Let $\lambda = \min_{i,j} \{-(d_j t_{ij} + r): (d_j t_{ij} + r) < 0\}$. If $\{i, j | (d_j t_{ij} + r) < 0\} = \emptyset$, let $\lambda = -r$. Let $r = r + \lambda$, and go to Step 1.

2.2 Let $r'$ be the point at which the two lines defined by $m_+$, $i_+$ and $m_-$, $i_-$ intersect. If $|r - r'| < \epsilon$, stop. Let $r = r'$, and go to Step 1.

Step (3): If $m > 0$, go to Step 3.1. Let $m_- = m$ and $i_- = i$. If $m_+ \neq 0$, go to Step 3.2. Let $\lambda = r$. Let $r = r - \lambda$, and go to Step 1.

3.1 Let $m_+ = m$ and $i_+ = i$. If $m_- \neq 0$, go to Step 3.2. Let $\lambda = \min_j \{d_j/(\bar{a}_j - a_j) | (\bar{a}_j - a_j) \neq 0\}$. Let $r = r + \lambda$, and go to Step 1.

3.2 Let $r'$ be the point at which the two lines defined by $m_+$, $i_+$ and $m_-$, $i_-$ intersect. If $|r - r'| < \epsilon$, stop. Let $r = r'$, and go to Step 1.

The updated (LD7) solution procedure was embedded in a branch-and-bound procedure similar to that provided earlier for (GBPP2). In fact, the only difference here is in the composition of Steps 0 and 1. We present the following specification accordingly:

*Algorithm for (GBPP2) with (LD7) Solution Incorporated*

Step (0): Let the subproblem list be empty, and let $p = 0$. Let the candidate bound $C^*$ be a large positive value. Perform Steps 1 through 4, and go to Step 5.

Step (1): Solve (LD7) to yield a current bound $C$ and a current solution $(x, y)$. If $C \geqslant C^*$, go to Step 4.

Steps (2), (3), (4), (5), and (6) are the same as for the previous (GBPP2) algorithm.

## 3. SAMPLE PROBLEM

In this section we present a small example problem in order to demonstrate the use of the approximate lagrangian formulation for (GBPP2). In particular, we present in step-by-step manner the application of the algorithm which concluded the previous section. Consider the problem given by the table below:

*Sample Problem:* $n = 5$, $m = 4$,

| $j$: | 1 | 2 | 3 | 4 |
|---|---|---|---|---|
| $c_j$: | 5 | 3 | 4 | 2 |
| $d_j$: | 2 | 1 | 3 | 2 |
| $w_j$: | 7 | 8 | 7 | 6 |

| $t_{ij}$: $i$ \ $j$ | 1 | 2 | 3 | 4 |
|---|---|---|---|---|
| 1 | 1 | 2 | 2 | 1 |
| 2 | 4 | 2 | 2 | 3 |
| 3 | 5 | 5 | 3 | 4 |
| 4 | 3 | 4 | 2 | 4 |
| 5 | 2 | 2 | 1 | 1 |

We would proceed as follows:

Step (0): $C^* = \infty$, $p = 0$.

Step (1): $r = 0$, $C = 5$, $y_2 = y_4 = 1$, $x_{12} = x_{22} = x_{42} = 1$
$x_{14} = x_{44} = x_{54} = 1$
$r = -2$, $C = 6$, $y_2 = y_4 = 1$, $x_{32} = 1$, $x_{44} = 1$
$r = -5/3$, $C = 6$, $y_2 = y_4 = 1$, $x_{12} = x_{32} = 1$
$x_{14} = x_{44} = x_{54} = 1$.

Step (2): Go to Step 3.

Step (3): $I_1 = \{1\}$, $I_2 = \{2\}$ $J_1 = \{2, 4\}$
$\theta_{14} = \max\{1/6, 1/3\} = 1/3$. Add the subproblem with $r = -5/3$, $C = 6$, and $x_{14}$ as the branching variable to the subproblem list.

Step (4): $p = 1$.

Step (5): Select the subproblem with $r = -5/3$, $C = 6$, and $x_{14}$ as the branching variable from the subproblem list.

Step (6): $x_{14} = 0$.

Step (1): $r = -5/3,\ C = 19/3,\ y_2 = y_4 = 1,\ x_{12} = x_{32} = 1$
$x_{44} = x_{54} = 1$
$r = 0,\quad C = 7,\ y_2 = y_4 = 1,\quad x_{12} = x_{22} = x_{42} = 1$
$x_{44} = x_{54} = 1.$

Step (2): Go To Step 3.

Step (3): $I_1 = \{4\},\ I_2 = \{3\},\ J_4 = \{2, 4\},$
$\theta_{44} = \max\{1, 2\} = 2.$
Add the subproblem with $r = 0,\ C = 7$, and $x_{44}$ as the branching variable to the subproblem list.

Step (4): $p = 2$
$x_{14} = 1, y_4 = 1.$

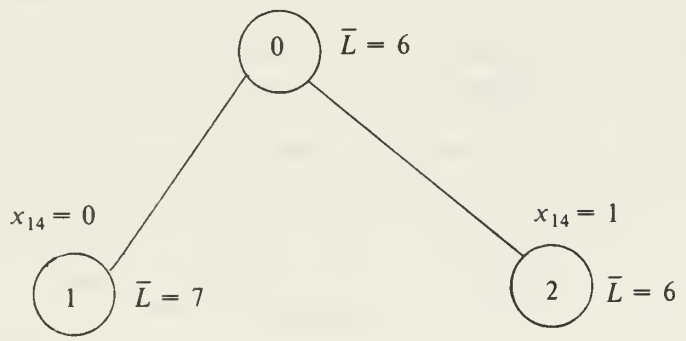
Step (1): $r = -5/3,\ C = 6,\ y_2 = y_4 = 1,\quad x_{22} = x_{32} = 1$
$x_{14} = x_{44} = x_{54} = 1.$

Step (2): $C^* = 6,\quad y_2^* = y_4^* = 1,\quad x_{22}^* = x_{32}^* = 1$
$x_{14}^* = x_{44}^* = x_{54}^* = 1.$
Remove the subproblem with $r = 0,\ C = 7$, and $x_{44}$ as the branching variable from the subproblem list.

Step (4): $p = 3.$

Step (5): Stop.

Schematically, the solution procedure can be summarized by the following tree:




Clearly, we have that chips 2 and 3 are packed into bin 2 at a total cost of 4 (bin cost of 3 plus unused penalty of 1) while chips 1, 4 and 5 are assigned to bin 4 at a cost of 2 (bin cost only, since capacity is used entirely) yielding the overall solution cost of 6.

## 4. COMPUTATIONAL RESULTS

In this section we present a summary of what was, in total, a rather substantive computational testing activity in this research. More precisely, we shall present empirical results relative only to the behavior of the approximate lagrangian models on randomly generated test problems. We remark, however, that additional tests were made for both the column generation and subgradient models in which the results were almost uniformly poorer than when the approximate lagrangian approach was examined. Hence, the stated concentration in this section follows accordingly.

To the extent that we present results for the case of only the approximate lagrangian models the following delineation must be made in order to properly interpret the ensuing tabular summaries. Recall that in the model given by (LD6) the dual multipliers specified by the vector $r$ were free in terms of sign distinction. Interestingly, if we restrict the sign of these multipliers such that $r \leqslant 0$, there is little effect in certain problem ranges, on overall computational effort. However, the mean number of dual iterations increases, thus indicating that the average price per dual iteration is less with sign restriction imposed than without such restriction. The following tables are organized around this differentiation.

In all tables, each of the parameters describing (GBPP1) are varied as dictated in the respective tables. In addition, for certain sizes, problems which were both correlated and uncorrelated (relative to $t_{ij}$) were generated and, in all experiments, two replicated at each level of $n$ and $m$ were performed. For ease of presentation, we have coded the ranges of parameter generation as follows:

$c_j \sim U(8, 10)$ _____ (a)
$c_j \sim U(40, 50)$ _____ (b)
$d_j \sim U(1, 2)$ _____ (c)
$d_j \sim U(5, 10)$ _____ (d)
$t_{ij} \sim U(4, 6)$ _____ (e)
$t_{ij} \sim U(1, 9)$ _____ (f)
$\rho = 0.0$ _____ (g)
$\rho = 0.5$ _____ (h).

Note that the space for ranges on the parameter $w_j$ is given directly in the tables since these ranges were generted in such a way as to preserve a constant average number of bins used in any solution. All computational experience reported was performed on a CYBER 74 computer.

In Table 1, a summary of the performance of the approximate lagrangian approach is presented for both the case of a restriction on the sign of the multiplier vector $r$ and the case where $r$ is not restricted. Calling these problems set $I$, we find that for cases $n = 10$, $m = 5, 7, 9$ and $n = 15$, $m = 5, 7, 9$ the procedure averaged less effort per problem where no restriction or $r$ was imposed. For the remaining levels of $n$ and $m$ (in Table 1) the opposite outcome occurred. Observing that problems at the larger values of $n$ dealt only with the "random" case of $\rho = 0.0$, there is obviously some cause for suspicion regarding the effect on the algorithm relative to "harder" versus "easier" problems. To lend some insight into this potential dichotomy in performance, a second problem set was run for those problems in Table 1 where the parametric range was fixed to coincide with those (the exception is with capacity) for the case $n = 25$ and $m = 5$. Note in Table 2, the cases of $n = 25$, $m = 5$ and $n = 20$, $m = 7$ are excluded since clearly, they would appear as in Table 1.

In observing the average computational effort required per problem solved in problem Set II, there does appear to indeed be at least empirical evidence that the restricted case out-performs the case when $r$ is subject to no sign restriction. Of course, if sufficient interest existed for making more solid any claims regarding the performances of both cases, a substantially greater experimental design than that used in this work would be in order. Regardless, to gain some insight into computational performance on larger problems we have applied only the restricted case to a third problem set. The results appear in Table 3.

An overall examination of the computational results reveals that sizeable effort can be required in a variety of instances. At the very least, there is substantial disparity in minimum and maximum requirements. Clearly, subtle parametric changes in problems otherwise identical can create vastly different computational demands.

TABLE 1 — *Approximate Lagrangian Procedure — Results for Problem Set I*

| Problem Size $n$ | $m$ | Problems Attempted Solved* | Parametric Ranges | Bin Capacity, $w_j$ | $r$ | Time† (Seconds) | Average Number of Nodes | Average Number of Dual Iterations |
|---|---|---|---|---|---|---|---|---|
| 10 | 5 | 64<br>64 | a,b,c,d,e,<br>f,g,h | $w_j \sim U(9, 19)$<br>$w_j \sim U(20, 30)$ | restricted | 2.780<br>(0.082, 26.998) | 91.531 | 101.625 |
| | | | | | unrestricted | 1.915<br>(0.079, 16.802) | 34.281 | 53.484 |
| 10 | 7 | 64<br>64 | a,b,c,d,e,<br>f,g,h | $w_j \sim U(9, 19)$<br>$w_j \sim U(20, 30)$ | restricted | 4.515<br>(0.058, 24.228) | 97.375 | 110.094 |
| | | | | | unrestricted | 3.740<br>(0.124, 17.463) | 52.250 | 75.563 |
| 10 | 9 | 64<br>64 | a,b,c,d,e,<br>f,g,h | $w_j \sim U(9, 19)$<br>$w_j \sim U(20, 30)$ | restricted | 8.580<br>(0.155, 82.342) | 174.656 | 184.750 |
| | | | | | unrestricted | 5.975<br>(0.422, 39.819) | 68.156 | 98.000 |
| 15 | 5 | 64<br>56 | a,b,c,d,e,<br>f,g,h | $w_j \sim U(16, 26)$<br>$w_j \sim U(33, 43)$ | restricted | 14.721<br>(0.529, 130.665) | 174.107 | 213.750 |
| | | | | | unrestricted | 12.707<br>(0.887, 118.463) | 66.071 | 109.107 |
| 15 | 7 | 32<br>31 | a,b,c,d,e,<br>f,g | $w_j \sim U(16, 26)$<br>$w_j \sim U(33, 43)$ | restricted | 12.917<br>(0.402, 95.639) | 157.452 | 190.097 |
| | | | | | unrestricted | 12.503<br>(0.839, 61.382) | 78.806 | 128.742 |
| 15 | 9 | 32<br>31 | a,b,c,d,e,<br>f,g | $w_j \sim U(16, 26)$<br>$w_j \sim U(33, 43)$ | restricted | 25.434<br>(1.040, 95.220) | 327.516 | 349.613 |
| | | | | | unrestricted | 20.590<br>(0.569, 85.658) | 87.516 | 159.097 |
| 20 | 5 | 32<br>24 | a,b,c,d,e,<br>f,g | $w_j \sim U(24, 34)$<br>$w_j \sim U(45, 55)$ | restricted | 22.775<br>(2.182, 93.355) | 129.000 | 170.013 |
| | | | | | unrestricted | 44.268<br>(5.571, 139.323) | 151.167 | 276.583 |
| 20 | 7 | 16<br>16 | a,b,c,d,e,<br>f,g | $w_j \sim U(45, 55)$ | restricted | 23.833<br>(1.476, 85.540) | 135.875 | 282.438 |
| | | | | | unrestricted | 41.482<br>(5.616, 121.813) | 112.500 | 186.438 |
| 25 | 5 | 16<br>9 | a,b,c,d,e,<br>f,g | $w_j \sim U(58, 68)$ | restricted | 14.372<br>(2.386, 59.413) | 53.222 | 81.111 |
| | | | | | unrestricted | 35.381<br>(15.762, 84.891) | 86.111 | 186.778 |

*A maximum of 1000 nodes in the branch-and-bound tree or 200 seconds of execution time (whichever was reached first) were used as cutoffs.

†Average (minimum, maximum)

TABLE 2 — *Approximate Lagrangian Procedure — Results for Problem Set II*

| Problem Size $n$ | $m$ | Problems Attempted Solved | Parametric Ranges | Bin Capacity, $w_j$ | $r$ | Time (Seconds) | Average Number of Nodes | Average Number of Dual Iterations |
|---|---|---|---|---|---|---|---|---|
| 10 | 5 | 16<br>16 | a,b,c,d,e,<br>f,g | $w_j \sim U(20, 30)$ | restricted | 0.937<br>(0.082, 4.697) | 30.250 | 32.063 |
| | | | | | unrestricted | 1.209<br>(0.201, 3.253) | 20.125 | 29.313 |
| 10 | 7 | 16<br>16 | a,b,c,d,e,<br>f,g | $w_j \sim U(20, 30)$ | restricted | 2.844<br>(0.190, 13.784) | 55.875 | 68.688 |
| | | | | | unrestricted | 3.690<br>(0.276, 17.463) | 42.750 | 77.750 |
| 10 | 9 | 16<br>16 | a,b,c,d,e,<br>f,g | $w_j \sim U(20, 30)$ | restricted | 4.409<br>(0.291, 38.743) | 85.250 | 93.625 |
| | | | | | unrestricted | 4.776<br>(0.422, 32.535) | 54.500 | 80.875 |
| 15 | 5 | 16<br>16 | a,b,c,d,e,<br>f,g | $w_j \sim U(33, 43)$ | restricted | 5.115<br>(0.620, 18.384) | 53.750 | 87.563 |
| | | | | | unrestricted | 8.936<br>(1.120, 45.392) | 46.375 | 95.688 |
| 15 | 7 | 16<br>16 | a,b,c,d,e,<br>f,g | $w_j \sim U(33, 43)$ | restricted | 7.770<br>(0.402, 22.011) | 76.875 | 121.813 |
| | | | | | unrestricted | 9.811<br>(1.390, 23.729) | 50.250 | 96.313 |
| 15 | 9 | 16<br>16 | a,b,c,d,e,<br>f,g | $w_j \sim U(33, 43)$ | restricted | 9.877<br>(1.424, 31.649) | 83.125 | 91.750 |
| | | | | | unrestricted | 13.860<br>(3.561, 74.410) | 21.875 | 44.438 |
| 20 | 5 | 16<br>15 | a,b,c,d,e,<br>f,g | $w_j \sim U(45, 55)$ | restricted | 20.226<br>(2.182, 63.068) | 78.600 | 117.067 |
| | | | | | unrestricted | 42.075<br>(5.571, 138.323) | 120.200 | 210.133 |

TABLE 3 — *Approximate Lagrangian Procedure—Results for Problem Set III (r restricted only)*

| Problem Size n | m | Problems Attempted Solved | Parametric Ranges | Bin Capacity, $w_j$ | Time (Seconds) | Average Number of Nodes | Average Number of Dual Iterations |
|---|---|---|---|---|---|---|---|
| 20 | 9 | 16<br>16 | a,b,c,d,e,<br>f,g | $w_j \sim$ U(45, 55) | 26.568<br>(3.402, 90.836) | 90.000 | 97.938 |
| 25 | 7 | 16<br>12 | a,b,c,d,e,<br>f,g | $w_j \sim$ U(58, 68) | 47.984<br>(3.546, 148.075) | 131.167 | 207.750 |
| 25 | 9 | 8<br>8 | a,b,c,d,f,<br>g | $w_j \sim$ U(58, 68) | 15.493<br>(6.420, 27.098) | 115.750 | 125.250 |
| 30 | 5 | 8<br>8 | a,b,c,d,f,<br>g | $w_j \sim$ U(70, 80) | 23.617<br>(14.446, 45.554) | 202.750 | 271.875 |
| 30 | 7 | 8<br>8 | a,b,c,d,f,<br>g | $w_j \sim$ U(70, 80) | 37.407<br>(10.793, 91.077) | 374.250 | 480.250 |
| 35 | 5 | 8<br>8 | a,b,c,d,f,<br>g | $w_j \sim$ U (83, 93) | 53.194<br>(10.667, 100.373) | 374.750 | 656.375 |
| 40 | 5 | 8<br>6 | a,b,c,d,f,<br>g | $w_j \sim$ U(95, 105) | 49.607<br>(31.770, 63.353) | 148.000 | 180.500 |
| 45 | 5 | 4<br>4 | a,b,d,f,g | $w_j \sim$ U(120, 130) | 92.894<br>(66.278, 125.692) | 364.500 | 426.250 |

Naturally, the performance of the procedure(s) tested here would be enhanced by direct comparisons with alternative approaches. The difficulty, however, is that such experience does not appear to be available relative to (GPBB1). Of course, various special cases have been examined as suggested earlier and to this extent the procedures tested here may be applied. This sort of comparative analysis would suffer, of course, since the attack presented here would carry unnecessary computational burdens not required in treating such special cases. In short, the models we have developed most fully and which have been tested in this analysis cater to the full interpretation of (GPBB1). Resolution of special cases, including the classic bin-packing problem, can be accomplished more efficiently by alternative means.

## 5. SUMMARY

Our entire concentration in this paper has dealt with the optimization of a combinational problem we have referred to as generalized bin-packing. Of course, one does not require the experience of this treatise to realize the difficulty of the problem; however, the empirical evidence gathered in this research does indeed tend to dismiss any suspicions to the contrary. In this light, a logical avenue would likely be to pursue some nonexact approaches to the problem in much the same way as has been done for the classic bin-packing model.

Of the various heuristic attacks we have postulated and examined, most were constructed directly from optimization procedures. Stated alternatively, specific heuristics involved suitable relaxation of the optimization models presented here. In other cases, nonexact schemes were developed as direct extensions of the any-fit heuristics discussed in Section 1. In general, procedures of the first category were not (empirically speaking) particularly noteworthy in that computational effort frequently rivaled that of the optimization attacks used. In the latter case, the more ad hoc procedures often generated solutions of almost arbitrarily poor quality resulting largely from an inability to make allowances for the unused capacity penalty.

It is worth adding that all testing of heuristic developments has been empirical in that no worst-case analysis was pursued. That this was the case is based not on a lack of interest per se but rather on the view that such results may be void of the sort of insight presented by similar results for the classic bin-packing problem. Again, the issues here would appear to relate to the arbitrary cost parameter structure of the problem.

## REFERENCES


[1] Baker, K., *Introduction to Sequencing and Scheduling,* (John Wiley and Sons, New York, 1974).

[2] Balas, E., "An Additive Algorithm for Solving Linear Programs with Zero-One Variables," Operations Research, *13*, 517-546 (1965).

[3] Coffman, E., Editor, *Computer and Job-Shop Scheduling Theory,* (John Wiley and Sons, New York, 1976).

[4] Coffman, E., M. Garey, and D. Johnson, "An Application of Bin Packing to Multi-Processor Scheduling," SIAM Journal of Computing, *7*, 1-17 (1978).

[5] De Maio, A. and C. Roveda, "An All Zero-One Algorithm for a Certain Class of Transportation Problems," Operations Research, *19*, 1406-1418 (1971).

[6] Eilon, S. and N. Christofides, "The Loading Problem," Management Science, *17*, 259-268 (1971).

[7] Eilon, S., C. Watson-Gandy, and N. Christofides, *Distribution Management* (Hafner Publishing Co., New York, 1971).

[8] Fisher, M., W. Northup, and J. Shapiro, "Using Duality to Solve Discrete Optimization Problems: Theory and Computational Experience," *Mathematical Programming Study 3* (North Holland Publishing Co., Amsterdam, 56-94, 1975).

[9] Fisk, J. and M. Hung, "A Heuristic Routine for Solving Large Loading Problems," Naval Research Logistics Quarterly, *26*, 643-650 (1979).

[10] Garey, M., R. Graham, and J. Ullman, "Worst-Case Analysis of Some Memory Allocation Algorithms," Proceedings of the Fourth Annual ACM Symposium on the Theory of Computing, 143-150, 1972.

[11] Garey, M., R. Graham, and J. Ullman, "An Analysis of Some Packing Algorithms," *Combinatorial Algorithms* (Academic Press, New York, 39-48, 1973).

[12] Garey, M. and D. Johnson, *Computers and Intractability: A Guide to the Theory of NP-Completeness* (W. H. Freeman and Co., San Francisco, 1979).

[13] Geoffrion, A. M., "Lagrangian Relaxation for Integer Programming," *Mathematical Programming Study 2* (North Holland Publishing Co., Amsterdam, 82-114, 1974).

[14] Gilmore, P. and R. Gomory, "A Linear Programming Approach to the Cutting Stock Problem: Part I," Opertions Research, *9*, 849-859 (1961).

[15] Gilmore, P. and R. Gomory, "A Linear Programming Approach to the Cutting Stock Problem: Part II," Operations Research, *11*, 863-888 (1963).

[16] Hung, M. and J. Brown, "An Algorithm for a Class of Loading Problems," Naval Research Logistics Quarterly, *25*, 289-297 (1978).

[17] Hung, M. and J. Fisk, "An Algorithm for 0-1 Multiple Knapsack Problems," Naval Research Logistics Quarterly, *25*, 571-579 (1978).

[18] Ingargiola, G. and J. Korsh, "An Algorithm for the Solution of 0-1 Loading Problems," Operations Research, *23*, 1110-1119 (1975).

[19] Johnson, D., "Near Optimal Bin Packing Algorithms," Doctoral Thesis, Massachusetts Institute of Technology, Cambridge (1973).

[20] Johnson, D., "Fast Algorithms for Bin Packing," Journal of Computer and Systems Science, *8*, 272-314 (1974).

[21] Johnson, D., A. Demers, J. Ullman, M. Garey, and R. Graham, "Worst-Case Performance Bounds for Simple One-Dimensional Packing Algorithms," SIAM Journal of Computing, *3*, 297-325 (1974).

[22] Lewis, R.T., "A Generalized Bin-Packing Problem," Doctoral Thesis, Georgia Institute of Technology, Atlanta (1980).

[23] Padberg, M., "On the Facial Structure of Set Packing Polyhedra," Mathematical Programming, *5*, 199-215 (1973).

[24] Ross, C. and R. Soland, "A Branch-and Bound-Algorithm for the Generalized Assignment Problem," Mathematical Programming, *8*, 91-103 (1975).

[25] Ross, G. and A. Zoltners, "Weighted Assignment Models and Their Application," Management Science, *25*, 683-696 (1979).

[26] Shapiro, J., "A Survey of Lagrangian Techniques for Discrete Optimization," Annals of Discrete Mathematics, *5*, 113-138 (1979).
[27] Srinivasan, V. and G. Thompson, "An Algorithm for Assigning Uses to Sources in a Special Class of Transportation Problems," Operations Research, *21*, 284-295 (1973).

# AN INTEGER PROGRAMMING ALGORITHM FOR PORTFOLIO SELECTION WITH FIXED CHARGES

Mary W. Cooper and Keyvan Farhangian

*Department of Operations Research*
*and Engineering Management*
*Southern Methodist University*
*Dallas, Texas*

ABSTRACT

A mean-variance portfolio selection model with limited diversification is formulated in which transaction and management costs are incorporated as the sum of a linear cost and a fixed cost. The problem is a fixed charge integer programming problem solved by hypersurface search using dynamic programming. Fathoming is performed in the forward pass of dynamic programming so that values of the state variable which correspond to infeasible solutions are eliminated from the tables. This logic permits the solution of problems with 20-30 possible investments.

## INTRODUCTION

This paper presents a portfolio selection model for the small investor. A basic reference for the portfolio selection problem is the book of Markowitz [4]. He presents the portfolio selection problem in which there are two criteria for the best portfolio: maximum return and minimum variance. Portfolios are called efficient if they meet either of the following conditions: if for a fixed variance the return is maximized, or if the return is fixed and the variance is a minimum. Sharpe [7] subsequently introduced some simplified models in which returns on securities are functions of the performance of one or more market indices. Recently Pang [6] has formulated and solved the portfolio analysis problem for large scale continuous variable problems by using a modified version of the parametric principle pivoting algorithm for the class of multiple index portfolio problems with positive covariance matrices. In the present paper we follow a different line of research. It has been shown by Jacob [3] that the solution of portfolio analysis problems by standard techniques generates answers which are inappropriate for the small investor because of the large number of positive amounts of different investments which they contain. The individual investor might prefer to limit the total *number* of his investments so that he does not have to pay extra charges on odd-lot sales and keep track of the performance of a great number of different securities. Jacob formulated several single index models of the portfolio analysis problem in which the total number of different investments is contrained. This results in an integer programming problem. Faaland solved another integer limited-diversification model parametrically to generate the set of efficient portfolios [2]. Our work extends his model by allowing transaction and management costs to be expressed explicitly as the sum of linear and fixed charges in the objective function. Then we use a technique for nonlinear integer programming which is based on dynamic programming [1].

### Notation for the Limited Diversification Model

Let us follow the notation for Sharpe's single index portfolio model. The return for security $j$ is $R_j$ and

$$R_j = E_j + \beta_j \{I - E(I)\} + \epsilon_j$$

where $E_j$ is expected return on security $j$, $I$ is the current value of the index and $E(I)$ is its expected value; $\beta_j$ is a parameter and $\epsilon_j$ is a random error term with mean zero which is independent of $I$. Define $\sigma^2(I)$ and $\sigma^2(\epsilon_j)$ as the variances of those variables and assume that $\text{cov}(\epsilon_i, \epsilon_j)$ is zero if $i \neq j$. Following Faaland, the integer problem in which the maximum variance $B_t$ is a parametric quantity is

$$\text{(1.1)} \quad \text{maximize} \quad \sum_{j=1}^{n} C_j X_j$$

$$\text{(1.2)} \quad \text{subject to} \quad \left[\sum_{j=1}^{n} a_j X_j\right]^2 + \sum_{j=1}^{n} Q_j X_j^2 \leqslant B_t$$

$$\text{(1.3)} \quad \sum_{j=1}^{n} X_j \leqslant k$$

(1.4) At most $T$ of the $x_j$ may be positive

$$\text{(1.5)} \quad x_j \geqslant 0 \qquad j = 1, \ldots, n$$

$$\text{(1.6)} \quad x_j \text{ integer} \qquad j = 1, \ldots, n$$

Let us define $C_j = E_j/k,\ j = 1, \ldots, n$

$$a_j = \beta_j\, \sigma(I)/k,\ j = 1, \ldots, n.$$

$$Q_j = \sigma^2(\epsilon_j)/k^2,\ j = 1, \ldots, n$$

$$B_t = \sigma_t^2,\ t = 1, \ldots, p.$$

The new problem includes a commission charge with a fixed minimum plus a component which is linear in the investment amount. Then (1.1) becomes

$$\text{(1.7)} \quad \text{maximize} \quad \sum_{j=1}^{n} (C_j - V_j)\, x_j - M_j \delta(x_j)$$

where $V_j$ is the linear charge, $M_j$ is the fixed charge, and

$$\delta(x_j) = \begin{cases} 1 & \text{if } x_j > 0 \\ 0 & \text{if } x_j = 0. \end{cases}$$

## THE INTEGER ALGORITHM

The problem described above (1.2)-(1.7) has several difficult aspects. In (1.2) the representation of portfolio variance is a nonseparable function. One method which can handle such functions is a nonlinear integer programming technique called hyperplane search by dynamic programming. Let us outline this approach. The general method is to solve a sequence of problems of the following form which differ only in the value of the right hand side by dynamic programming.

$$\text{maximize} \quad z = \Sigma\, f_j(x_j)$$

$$\Sigma\, f_j(x_j) = z_k$$

$$0 \leqslant l_j \leqslant x_j \leqslant u_j \quad j = 1, \ldots, n$$

$$x_j \text{ integer} \qquad j = 1, \ldots, n.$$

This problem can be used to identify all integer points on any hypersurface level $z_k$. If $z_k$ is an upper bound on the objective function value, then the solution of the single dynamic programming problem above can be used to generate all integer points upon the hypersurface level $\Sigma f_j = z_k$ and for all right hand side values less than $z_k$. In fact it gives information as to which values of $z$ contain integer points so that hypersurfaces with no such points need never be considered, see [1]. We can use this technique on (1.7). Then the constraints of the original problem (1.2)-(1.6) are used to test the feasibility of any candidate integer points, so that the intractability of (1.2) can be overcome. If we choose the sequence $\{z_k\}$ in decreasing order (or generate integer points on hypersurface levels with decreasing objective function values) then the first feasible point is optimal. If $\{z_k\}$ is decreasing then this method is a partial enumeration since no integer points are considered which have objective function values less than the optimal solution. This strategy is a direct application of the hyperplane search method to limited diversification portfolio analysis. The complicated form of the objective function (1.7) presents no difficulty.

**A Second Approach**

Since the straightforward method given above requires too much storage, the usual method of calculating the dynamic programming recursion tables ("pulling") was changed to a second approach which uses a technique called "pruning" or "fathoming" to reduce the size of the tables by allowing infeasible points to be eliminated during the forward pass. In this way, the constraints (1.2)-(1.4) can be used to reduce computation and storage. The idea assumes that coefficients in the constraints are positive. Under this assumption, if we have a partial solution for which the portfolio variance constraint (1.2) (or (1.3) or (1.4)) is exceeded, then we "prune" the entry or value of element $x_s$ which caused the infeasibility since no completion of that partial solution will be feasible [5]. This method of fathoming reduces both the storage and computational burden of the algorithm. It supplants the method of checking partial solutions for infeasibility at every stage in the backward pass of dynamic programming.

**Implementation of the Parametric Right-Hand-Side Feature**

The problem (1.2)-(1.7) has a parametric right hand side so that efficient portfolios are generated for various upper bounds on portfolio variance. Previous work [1] has shown that such parametric problems can be solved from a single dynamic programming problem using hypersurface search, because of a similar property obtained in any dynamic programming solution. If an allocation problem is solved for a given right-hand side value, then by completing the last return function table for all values of the state variable all problems of identical form but with smaller right hand side values can be obtained.

**Computational Results**

The model selected for computational results is that given in (1.2)-(1.7), a nonlinear, nonseparable parametric fixed charge problem. Problem coefficients were generated randomly for various numbers of total possible investments from 10-30. The times for computer runs on the CYBER 73 are given in the following table of summary statistics. For each problem size, statistics are averaged for ten problems with random coefficients. The second through the tenth problems are parametric versions of the first and therefore take less time.

This problem is highly specialized. Therefore it is difficult to find computational comparisons with it in the literature. However some recent work on a linear integer bicriteria problem [8] may be relevant. Villarreal and Karwan have given results for three ten-variable problems with four constraints which are solved in an average of 18.43 CPU seconds each on a CDC 6400 computer system. Their algorithm generates all distinct efficient points (an average of 7) in comparison to ours which generates efficient points according to ten desired levels of portfolio variance. These authors obtain additional

TABLE 1 — *($\mu$/$\sigma$ in CPU Seconds)*

| Portfolio Variance = $B_t$ | $n = 10$ | $n = 20$ | $n = 30$ |
|---|---|---|---|
| 1.5 | 0.1/.06 | 6.2/7.3 | 34.74/36.43 |
| 1.35 | .00/.00 | .01/.01 | .01/.02 |
| 1.2 | .01/.01 | .08/.19 | .08/.15 |
| 1.05 | .00/.01 | .04/.09 | .54/.85 |
| .9 | .00/.01 | .28/.49 | .19/.44 |
| .75 | .01/.02 | .43/.64 | 2.59/4.62 |
| .6 | .01/.01 | 1.12/2.06 | .81/1.20 |
| .45 | .01/.02 | 1.04/1.68 | 8.78/11.47 |
| .3 | .02/.03 | 2.38/4.29 | 3.64/6.85 |
| .15 | .02/.02 | 2.36/3.16 | 11.96/13.32 |

improvements with various hybrid dynamic programming recursions, and get their best times (1.19 CPU seconds for ten variables) with a branch and bound approach. Our method seems competitive, particularly when the nonlinearities and nonseparability of the portfolio problem are considered, along with the larger size of problems solved. It is difficult to make a more exact comparison because of the different structure of the problems in (8) and the portfolio problem.

## REFERENCES

[1] Cooper, M., "Postoptimality Analysis in Nonlinear Integer Programming," Naval Research Logistics Quarterly, *28*, 301-307 (1981).

[2] Faaland, B., "An Integer Programming Algorithm for Portfolio Selection," Management Science, *20*, 1376-1384 (1974).

[3] Jacob, N., "A Limited-Diversification Portfolio Selection Model for the Small Investor," Journal of Finance, *29*, 847-852 (1974).

[4] Markowitz, H., Portfolio Selection: Efficient Diversification of Investments (John Wiley and Sons, New York, 1959).

[5] Morin, T., "Computational Advances in Dynamic Programming," Dynamic Programming and Its Applications, M. Puterman, Editor (Academic Press, New York, 1978).

[6] Pang, J., I. Kaneko and W. Hallman, "On the Solution of Some (Parametric) Linear Complementarity Problems with Applications to Portfolio Selection, Structural Engineering and Acturial Gradation," Mathematical Programming, *16*, 325-347 (1979).

[7] Sharpe, W., "A Linear Programming Algorithm for Mutual Fund Portfolio Selection, Management Science, *13*, 499-510 (1967).

[8] Villarreal, B. and M.H. Karwan, "Multicriteria Integer Programming: A (Hybrid) Dynamic Programming Recursive Approach," Mathematical Programming, *21*, 204-223 (1981).

# SCHEDULING OF A SINGLE MACHINE TO MINIMIZE TOTAL WEIGHTED COMPLETION TIME SUBJECT TO RELEASE DATES

Lucio Bianco and Salvatore Ricciardelli

*Istituto di Analisi dei Sistemi ed Informatica del CNR*
*Rome, Italy*

ABSTRACT

In this paper the $n/1/r_j \geqslant 0/\sum_j w_j C_j$ problem under the assumptions of nonpreemptive sequencing and sequence independent processing times is investigated. After pointing out the fundamental properties, some dominance sufficient conditions among sequences are obtained and a branch and bound algorithm is proposed. Computational results are reported and discussed.

## 1. INTRODUCTION

The one machine scheduling problem has been studied extensively under different hypotheses and objective functions. Nevertheless, in the literature, emphasis is laid upon the case with equal ready times, no imposed due dates and flow time (or equivalent function) as objective function. In this context the problem solution is trivial since the Smith's rules [14], well-known as shortest processing time (SPT) rule and weighted shortest processing time (WSPT) rule, provide an optimal solution. According to these rules, jobs are sequenced on the basis of a preestablished order of their processing times.

In [7], [8], [12], [13], and [15], a more sophisticated cost function, such as weighted tardiness, is adopted for it turns out to be a suitable model for many real problems in which jobs may be considered available simultaneously for processing on a single machine in order to minimize the total sum of the delays with respect to their due dates. One of the most constraining assumptions is equality of the ready times of the jobs.

As a matter of fact, whenever different ready times are to be considered ($r_j \geqslant 0$), the previous models and the corresponding algorithms are no more adequate.

The case $r_j \geqslant 0$ has been recognized in research on other single machine problems, where due dates were also taken into consideration [1], [5], and [10]. However, in these works, the authors investigate the properties of a rather different cost function given by the maximum lateness (or tardiness) with weights $w_j = 1$ for all $j$.

In [2], [3], [4], and [11], referring to scheduling theory application to air traffic control, minimizing total weighted waiting time in one machine scheduling problem with unequal ready times has been considered.

An implicit enumeration procedure has been developed and a simplified algorithm ($w_j = 1, \forall j$) has been implemented in order to work in real time. The interested reader may address to references [11] and [4] for a complete discussion of this application. The aforementioned scheduling problem, although simple in these terms, is not trivial and it may be proved to be NP-complete [9]. The main difficulty arises from the fact that, since $r_j \geqslant 0$, idle times may be inserted in the optimal schedule [6].

In this paper, starting from the results obtained in previous works of ours [2], [3], [4], and [11] the $n/1/r_j \geqslant 0/\sum_j w_j C_j$ problem, with no preemption allowed and sequence independent processing times, is analyzed. The purpose is to establish some further dominance properties and above all to present a much improved branch and bound algorithm. Computational results, obtained with this algorithm, are also reported and discussed.

## 2. DEFINITIONS AND PROBLEM STATEMENT

Let $N = \{j | j = 1, 2, \ldots, n\}$ be a set of jobs to be processed, one job at a time, on a single, continuously available, machine. For each job $j$, the ready time $r_j$, the processing time $p_j$, and the weight $w_j$, are given.

We call sequence on set $K \subset N$ any permutation in $K$ which is indicated with $s_k$, being $s_0$ the sequence on the void set. Completion of all jobs requires establishing a sequence $s_n = (j_{[1]}, j_{[2]}, \ldots, j_{[n]})$. When a job parameter is identified by the job's position in a given sequence rather than its index number, the position is indicated in square brackets. Thus $C_{[j]}$ means the completion time of whichever job occupies the $j$th position in the sequence.

Suppose now that a sequence is constructed by adding one job at a time, starting from position 1. At any step, we have a partial sequence $s_k = (j_{[1]}, j_{[2]}, \ldots, j_{[k]})$ of a job set $K \subset N$. It can easily be verified that, in our setting, the completion time of such a sequence cannot be expressed only in terms of processing times regardless of the order in which jobs are scheduled, since idle times can be inserted [6]. This fact suggests some useful definitions.

DEFINITION 1: Given a partial sequence $s_k$, the earliest start time of a job $j \in \bar{K}$, where $\bar{K} = N - K$ is the set of jobs not sequenced, can be expressed as

$$t_j(s_k) = \max(r_j, C_{[k]}) \tag{1}$$

with $t_j = r_j$, $\forall j \in N$, if $k = 0$.

DEFINITION 2: Given a partial sequence $s_k$, the completion time of a job $j \in \bar{K}$, $C_j(s_k)$, can be expressed as

$$C_j(s_k) = t_j(s_k) + p_j. \tag{2}$$

DEFINITION 3: Let $K$ be any subset of $N$ with cardinality $k$, then the cost function associated to $s_k$ is

$$C^*(s_k) = \sum_{j=1}^{k} w_{[j]} C_{[j]}.$$

If $k = n$, $C^*(s_n)$ expresses the total weighted completion time of $s_n$.

Then, the problem can be formally stated as follows: Given a job set $N$, find a sequence $s_n^*$ such that

$$C^*(s_n^*) \leqslant C^*(s_n), \ \forall s_n \in S_N.$$

To fully understand the problem framework, it is suitable to introduce the reduced problem concept. A definition of reduced problem may be found in references [7] and [15]. However, that definition is given by considering the sequence completion time as computed regardless of the order of the jobs. As pointed out before, this is impossible in our setting; a more general definition, which also holds in our hypotheses, is the following.

DEFINITION 4: A problem is said to be reducible if a partition $(N_1, N_2)$ of set $N$ exists such that, if $s_n^*$, $s_1^*$, $s_2^*$ are the optimal sequences with respect to $C^w(\cdot)$ on $N$, $N_1$, $N_2$, respectively, it is true that $s_n^* = s_1^* s_2^*$.

REMARK 1: It is easy to prove that a given problem $n/1/r_i \geqslant 0/\sum_j w_j C_j$ is certainly reducible if a job set $N_1$ and a corresponding optimal sequence $s_1$ exist such that completion time of $s_1$ is less than (or at most equal to) the earliest ready time associated with $N_2$.

In [11] a simple decomposition method to transform a given problem into reduced problems is outlined.

Therefore, without loss of generality, from now on, only reduced problems will be dealt with.

DEFINITION 5: Given a partial sequence $s_k$, a job set $V \subset \bar{K}$ is said to be "dense" if any sequence $j_1, j_2, \ldots, j_m$ on $V$ has no inserted idle times.

It is easy to verify that

$$j \in V \text{ iff } t_j(s_k) < \min_{j \in \bar{K}} C_j(s_k).$$

DEFINITION 6: Given two sequences $s_k$ and $s_h$ both belonging to set $S$, $s_k$ dominates $s_h$, if and only if, it is true that there exists an $s_{n-k}$ such that $s_k s_{n-k} = s_n'$ belongs to $S_N$ and

$$C^w(s_n') \leqslant C^w(s_n|s_h) \text{ for all } s_n|s_h.$$

DEFINITION 7: Two sequences, $s_k$ and $s_h$, are equivalent

$$\text{if } C^w(s_n^*|s_k) = C^w(s_n^*|s_h).$$

REMARK 2: On the basis of the previous definitions it follows immediately that an optimal sequence $s_n^*$ dominates any other one, and it is equivalent to all its own partial sequences $s_k$, $k = 1, 2, \ldots, n$.

DEFINITION 8: A sequence is said to be an ECT sequence if it satisfies the earliest completion time (ECT) rule, i.e., if $C_{[k+1]} = \min_{j \in \bar{K}} C_j(s_k)$, $0 \leqslant k \leqslant n-1$.

Ties are broken by choosing $j$ with min $t_j$ and further ties by choosing $j$ with max $w_j$ and, at last, by choosing $j$ with min $j$.

DEFINITION 9: A sequence is said to be an EST sequence if it satisfies the earliest start time (EST) rule, i.e., if each actual job start time is such as

$$T_{[k+1]} = \min_{j \in \bar{K}} t_j(s_k), \quad 0 \leqslant k \leqslant n-1.$$

## 3. OPTIMALITY AND DOMINANCE CRITERIA

In this section a set of properties, which enhance the efficiency of the search for an optimal solution, is given. Some of these properties (specifically those resulting from Theorems 2, 3, 4, 5 and Corollary 2) have already been demonstrated in a previous paper [11]. The other properties, expressed as Theorems 1, 6, 7 and Corollaries 1 and 3, have here been carried out in order to improve the algorithm performance. Extended proofs of these further properties are reported in the Appendix.

THEOREM 1: Given a job set $N$ and a partial sequence $s_k (k < n)$, if, for $i \in \bar{K}$

(a) $\frac{p_i}{w_i} \leqslant \frac{p_j}{w_j}$, $\forall j \in \bar{K}$

(b) $t_i(s_k) \leqslant t_h(s_k)$, $h \in \bar{K}$

then $s_k i$ "dominates" $s_k h$.

Clearly, the property expressed in Theorem 1, in the case with $w_i = 1 \ \forall i$, can be applied in the following way: "Given a partial sequence $s_k$, consider, among all the remaining jobs, job $i$ with the least processing time and then take out of consideration for $k + 1$th position, all jobs unable to start before job $i$. This property, of course, appears to be efficient since earlier job $i$ is able to start. For instance, if job $i$ has both minimum processing time and earliest start time, no other job may be considered for possible inclusion in the $(k + 1)$th position in the sequence.

COROLLARY 1: An EST sequence of job set $N$ is optimal if it is ordered according to WSPT rule.

THEOREM 2: Given a job set $N$, a partial sequence $s_k$ $(k < n)$ and two jobs $i,j \in \bar{K}$. If $r_j \geqslant C_i(s_k)$, then $s_k i$ dominates $s_k j$.

From this theorem it follows that is is possible to restrict search for optimal solution to the class of active schedules [1].

In Figure 1 an example of the application of Theorem 2 is given.

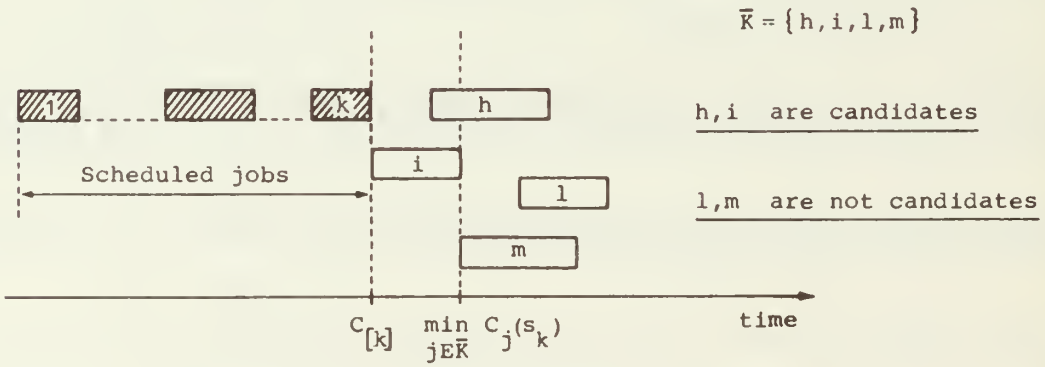



FIGURE 1. Example of application of Theorem 2: candidates to $(k + 1)$th position in the sequence.

THEOREM 3: Given $s_k$ and two jobs $i, j \in \bar{K}$, if

(a) $w_i \geqslant w_j$

(b) $C_j(s_k) \geqslant C_i(s_k)$

(c) $w_i[C_j(s_k) - C_i(s_k)] + w_i p_i - w_j p_j \geqslant (p_j - p_i) \sum_{l \in Q} w_l$

where $Q$ is the set $\bar{K} - \{i, j\}$, then $s_k i$ dominates $s_k j$.

THEOREM 4: Given $s_k$ and two jobs $i, j \in \bar{K}$, if

(a) $w_i \geqslant w_j$

(b) $C_j(s_k) \leqslant C_i(s_k)$

(c) $w_i p_i - w_j p_j \geqslant [C_i(s_k) - C_j(s_k)] \sum_{l \in \bar{K}} w_l + \delta_{-1}(p_j - p_i) \cdot \left[ \sum_{l \in \bar{K}} w_l - w_j - w_i \right]$

then $s_k i$ dominates $s_k j$.

Properties resulting from Theorems 3 and 4 may be practically utilized in order to operate a further selection within the class of active schedules. Of course, from a theoretical point of view they could also be applied within the more general class of permutation schedules. It is easily seen that, in general, the dominance concept, as expressed in Theorems 3 and 4, depends not only on the input variables (weight coefficients, processing and ready times) associated to the couple of jobs taken into consideration but also on the input variables associated to the remaining jobs (specifically on their weight coefficients or on their number if $w_i = 1 \; \forall i$).

THEOREM 5: Given a job set $N$ and two partial sequences $xiyj$ and $xjyi$. If

(a) $iyj$ and $jyi$ have no inserted idle times

(b) $[C_j(x) - C_i(x)] \sum_{l \in P} w_l + [w_i - w_j] \sum_{l \in P} p_l + w_j p_i - w_i p_j \geqslant \delta_{-1} [t_i(x) - t_j(x)] \sum_{l \in U} w_l,$

where $P$ is the job set of $iyj$ and $U$ the set $N - P - \{x\}$, then $xiyj$ dominates $xjyi$.

COROLLARY 2: Given $s_k$ and two jobs $i, j \in \bar{K}$. Let $\bar{K}$ be the set "dense," if

(a) $w_i \geqslant w_j$

(b) $t_j(s_k) \geqslant t_i(s_k)$

(c) $C_j(s_k) \geqslant C_i(s_k)$

then $s_k i$ dominates $s_k j$.

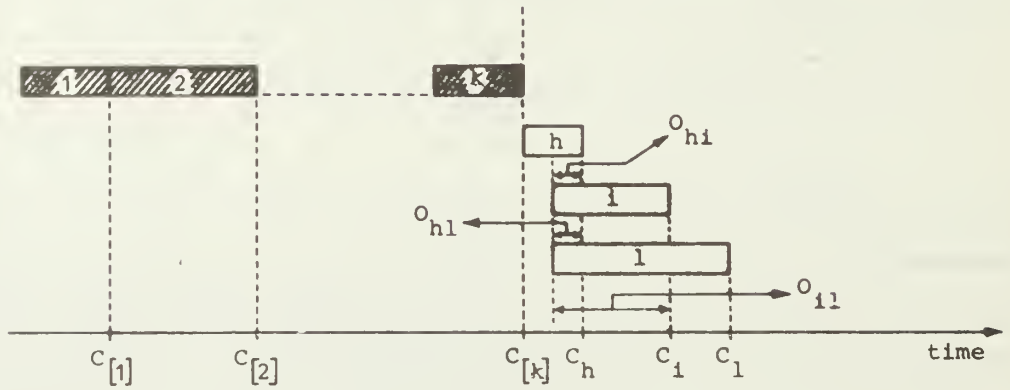
THEOREM 6: Given a job set $N$ and a partial sequence $s_k (k < n)$, for any sequence $(s_n | s_k)$ a lower bound $\underline{C}^w(s_n | s_k) \leqslant C^w(s_n | s_k)$ exists, where

$$\underline{C}^w(s_n | s_k) = C^w(s_k) + \sum_{j=1}^{n-k} w_{k+j} C_{k+j}(s_k) + \sum_{j=1}^{n-k-1} \sum_{l=i+1}^{n-k} \{\delta_{-1}[C_{k+j}(s_k) - t_{k+l}(s_k)] - \delta_{-1}[C_{k+j}(s_k) - C_{k+l}(s_k)]\} \cdot \min(w_{k+j}, w_{k+l}),$$

being $k + j$ and $k + l \in \bar{K}$ with $r_{k+j} \leqslant r_{k+l}$.

The first term of the lower bound provided by Theorem 6 is, simply, the value of the objective function associated with the partial sequence $s_k$. The second term takes into account that each job of the complementary set $\bar{K}$ cannot start before its earliest start time. The third term takes into account that processing times of the unscheduled jobs are not to be overlapped in the final sequence.

An example of application of this lower bound is reported in Figure 2.




$$\underline{C}^W(s_n/s_k) = W_{[1]}C_{[1]} + \cdots W_{[k]}C_{[k]} + W_hC_h + W_iC_i + W_lC_l + O_{hi}\cdot\min(W_i,W_h)$$

$$+ O_{hl}\cdot\min(W_h,W_l) + O_{il}\cdot\min(W_i,W_l)$$

FIGURE 2. Example of lower bound computation for a partial sequence $s_k$.

THEOREM 7: Let $s_n^0$ be a sequence on $N$ such that, for $0 \leqslant k \leqslant n-1$

(a) $T_{[k+1]} = C_{[k]}$

(b) $C_{[k+1]} = \min_{j \in \bar{K}} C_j(s_k)$

(c) $w_{[k]} \geqslant w_{[k+1]}$

being $C_{[0]} = r_i$, $i \in N$, then $s_n^0$ dominates any other sequence $s_n^a$ of the same set starting with job $h$ such as $r_i \leqslant r_h$.

COROLLARY 3: An ECT sequence, defined on a job set $N$ and ordered according to nondecreasing weights, is optimal if it has no inserted idle times and begins with the job having the smallest ready time.

## 4. THE PROPOSED APPROACH

To solve the problem under study, an implicit enumerative procedure is proposed, based upon a branch and bound concept. It involves a search along the branches of a tree, in which any node at level $k$ represents a partial sequence $s_k$ defined on a set $K$ with cardinality $k$.

As a consequence, subsequent jobs in $s_k$ correspond also to a succession of nodes encountered by going from tree root $s_0$ to node $s_k$. For each node $s_k$, as illustrated in Section 3, a lower bound $\underline{C}^w(s_n|s_k)$ on $C^w(s_n|s_k)$ is computed, and the current optimal solution is utilized as upper bound. At each iteration, the node being expanded is called the current node, whose last job has the lowest current value of earliest start time. A closed node is one whose correspondent partial sequence has been dominated and, therefore, it will be no longer considered. To identify dominated nodes, some elimination conditions based upon the previous theorems are then applied. In this way, the number of nodes branching from the current node can be effectively reduced.

The proposed procedure is a recursive one. In fact, whenever a partial sequence $s_k$ on $N$ has been fixed, the optimal sequencing problem of the jobs belonging to set $\bar{K}$ is identical to the initial problem if their ready times are shifted so that none of them is less than $C(s_k)$.

## 5. THE ALGORITHM

The algorithm consists of three basic phases: initialization, branching and termination. Initialization involves defining initial value of the variables, determining an initial sequence and testing it for optimality.

Branching is an iterative procedure. In each iteration, it generates the descendents of the current node, eliminates closed nodes, updates the set of job candidates to successive position, identifies, as next current node, that one having last job with lowest earliest start time, and computes the lower bound for it.

The termination phase consists of identification of the optimal sequence $s^*$ and computation of $C^w(s^*)$.

### Initialization

(1) Let $N$ be the job set $\{j | j = 1, 2, \ldots, n\}$. Define the corresponding parameter sets:

$$R = \{r_1, r_2, \ldots, r_j, \ldots, r_n\},\ P = \{p_1, p_2, \ldots, p_n\},$$

$$W = \{w_1, w_2, \ldots, W_n\} \text{ where } r_{j-1} \leqslant r_j \leqslant r_{j+1}.$$

If $r_j = r_{j+1}$ call $j$ job with min $p_j$

(2) $k = 0,\ K = \varnothing,\ s_k = \varnothing,\ \bar{K} = N,\ C(s_k) = r_1.$

(3) Construct an initial solution $\tilde{s}_n$ by means of an heuristic rule (for example EST or ECT) and compute $C^w(\tilde{s}_n)$.

(4) Test initial solution for optimality by means of Corollary 1 and Corollary 3. If $\tilde{s}_n$ passes the test go to Step 22.

### Branching

(5) Determine $t_j(s_k) = \max_{j \in \bar{K}} \{r_j, C(s_k)\}$ and

$$C_m(s_k) = \min_{j \in \bar{K}} \{t_j(s_k) + p_j\}$$

(6) Form the set

$$D_{s_k} = \{j | j \in \bar{K}, t_j(s_k) < C_m(s_k)\}$$

(7) If $D_{s_k} = \bar{K}$ go to Step 18

(8) Generate nodes $s_k j,\ \forall j \in D_{s_k}$

(9) For each node $s_k j$ dominated according to some of the Theorems 1, 3, 4, 5 (and Corollary 2 (if Step 8 has been reached coming from Step 19)), update set $D_{s_k}$ by setting

$$D_{s_k} = D_{s_k} - j$$

(10) Set $k = k + 1$

(11) If $D_{s_{k-1}} = \varnothing$, set $k = k - 1$ and go to Step 16

(12) Set $D_{s_{k-1}} = D_{s_{k-1}} - j'$, $s_k = s_{k-1} j'$, $K = \{s_k\}$, $\bar{K} = N - K$ with $j'$: $t_{j'} = \min_{j \in D_{s_{k-1}}} t_j(s_{k-1})$

(13) Compute lower bound $\underline{C}^w(s_n|s_k)$

(14) If $\underline{C}^w(s_n|s_k) \geqslant C^w(\tilde{s}_n)$ go to Step 11.

(15) Set $C(s_k) = t_{j'} + p_{j'}$ and go to Step 5.

(16) If $k = 0$, go to Step 22.

(17) Go to Step 11.

(18) If $t_j(s_k) = t, \forall j \in \bar{K}$, go to Step 20.

(19) Go to Step 8.

(20) Form the sequence $s_k s_{n-k}$, where $s_{n-k}$ is the sequence on $\bar{K}$ satisfying the WSPT rule.

(21) Set $\tilde{s}_n = s_k s_{n-k}$, $C^w(\tilde{s}_n) = \underline{C}^w(s_n|s_k)$ and go to Step 11).

**Termination**

(22) Set $s_n^* = \tilde{s}_n$ and $C^w(s_n^*) = C^w(\tilde{s}_n)$. STOP.

Figure 3 shows a synthetic flow-chart which illustrates the algorithm logic.

## 6. COMPUTATIONAL RESULTS AND CONCLUSIONS

The algorithm, previously illustrated, has been implemented on a Digital PDP 11/34 minicomputer.

Different series of tests have been performed by considering a given number of jobs $n = 10$. For each series the $p_i$ values have been chosen uniformly distributed over a fixed interval $[1 \div 7]$ and the $r_i$ values have been chosen uniformly distributed over five different time intervals of growing length. Moreover, for each fixed range of $r_j$, the weights have also been taken as uniformly distributed over ten different intervals of growing length. The results are fully reported in Table 1 while in Figures 4 and 5 are reported the diagrams showing the number of nodes generated versus max $r_j$ and max $w_j$, respectively. As it can be observed a significant increase in all the parameters is met whenever weight coefficients tend to be unequal. In particular, computational results show that changes in the observed parameters are relatively less dependent on the width of the weights range. Rather, these changes seem to arise whenever a mere difference among the weights is considered. Figure 4 and 5 show this behavior for the number of nodes generated.

Nevertheless, the computation time, in the worst case, does not exceed a few tens of seconds, even though a fast computer has not been used.

A last remark, involves the fact that the typical behavior (bell-shaped), observed in [4] when $w_j = 1, \forall j$, is confirmed also if $w_j$ are distributed on a fixed interval (Figure 4).

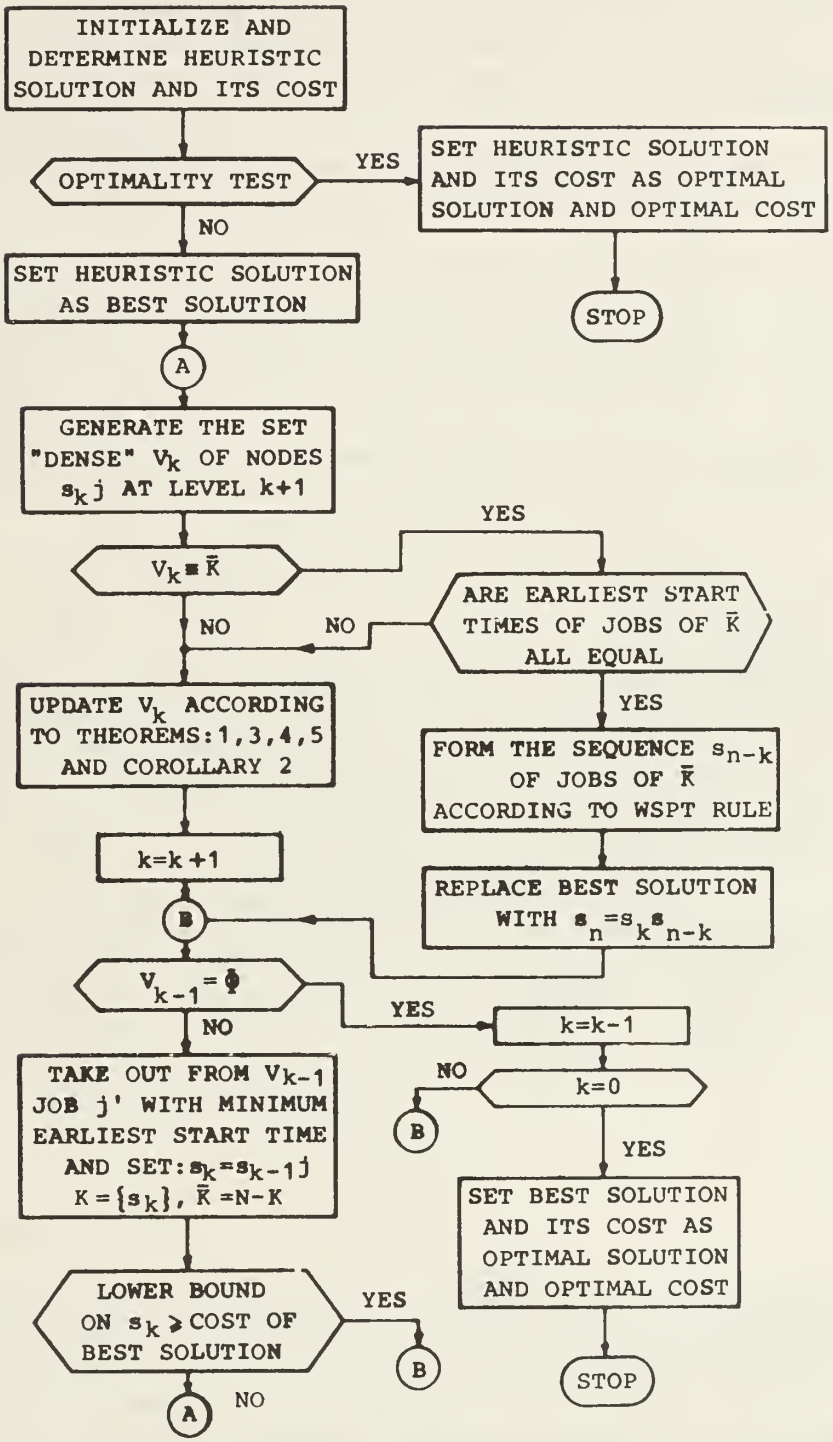

INITIALIZE AND DETERMINE HEURISTIC SOLUTION AND ITS COST
OPTIMALITY TEST
YES
SET HEURISTIC SOLUTION AND ITS COST AS OPTIMAL SOLUTION AND OPTIMAL COST
STOP
NO
SET HEURISTIC SOLUTION AS BEST SOLUTION
A
GENERATE THE SET "DENSE" $V_k$ OF NODES $s_k j$ AT LEVEL k+1
YES
$V_k \equiv \bar{K}$
ARE EARLIEST START TIMES OF JOBS OF $\bar{K}$ ALL EQUAL
NO
NO
YES
UPDATE $V_k$ ACCORDING TO THEOREMS: 1, 3, 4, 5 AND COROLLARY 2
FORM THE SEQUENCE $s_{n-k}$ OF JOBS OF $\bar{K}$ ACCORDING TO WSPT RULE
k=k+1
REPLACE BEST SOLUTION WITH $s_n = s_k s_{n-k}$
B
$V_{k-1} = \phi$
YES
k=k-1
NO
TAKE OUT FROM $V_{k-1}$ JOB j' WITH MINIMUM EARLIEST START TIME AND SET: $s_k = s_{k-1} j$ $K = \{s_k\}$, $\bar{K} = N - K$
NO
k=0
B
YES
SET BEST SOLUTION AND ITS COST AS OPTIMAL SOLUTION AND OPTIMAL COST
LOWER BOUND ON $s_k \geq$ COST OF BEST SOLUTION
YES
B
STOP
A
NO

FIGURE 3. Flow chart of the algorithm.

TABLE 1 — *Summary of the Results in the Case $n = 10$ and Range of $p_j = 1 \div 7$. Number of Tests for Each Pair of Intervals Relative to $r_j$ and $w_j = 100$. Total Number of Tests Performed = 5000*

| Range of $r_j$ / Range of $w_j$ | 1 ÷ 20 | | | 1 ÷ 30 | | | 1 ÷ 40 | | | 1 ÷ 50 | | | 1 ÷ 60 | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | a | b | c | a | b | c | a | b | c | a | b | c | a | b | c |
| 1 ÷ 1 | 27 | 18 | 14.5 | 30 | 26 | 12.8 | 24 | 22 | 9.5 | 16 | 13 | 7.5 | 15 | 15 | 6.8 |
| 1 ÷ 1.2 | 44 | 25 | 22.2 | 42 | 29 | 17.7 | 35 | 24 | 12.3 | 28 | 22 | 9.6 | 24 | 20 | 8.6 |
| 1 ÷ 1.4 | 42 | 24 | 22.1 | 46 | 32 | 19. | 40 | 28 | 13.3 | 31 | 22 | 9.8 | 25 | 20 | 8.6 |
| 1 ÷ 1.6 | 44 | 25 | 22.9 | 49 | 33 | 19.4 | 41 | 32 | 14.1 | 29 | 22 | 10 | 25 | 19 | 8.8 |
| 1 ÷ 1.8 | 44 | 25 | 23.8 | 49 | 30 | 19.4 | 45 | 32 | 15 | 29 | 25 | 9.9 | 24 | 20 | 8.8 |
| 1 ÷ 2.0 | 47 | 27 | 25 | 48 | 34 | 20.3 | 44 | 34 | 14.6 | 30 | 23 | 9.9 | 26 | 20 | 8.8 |
| 1 ÷ 2.2 | 47 | 26 | 25.5 | 55 | 39 | 22 | 41 | 34 | 14.9 | 31 | 25 | 9.8 | 26 | 21 | 8.9 |
| 1 ÷ 2.4 | 46 | 27 | 25.1 | 48 | 32 | 20.2 | 43 | 29 | 14.8 | 28 | 23 | 10.1 | 26 | 21 | 8.7 |
| 1 ÷ 2.6 | 48 | 26 | 25.1 | 52 | 37 | 21.4 | 40 | 31 | 13.9 | 27 | 22 | 9.7 | 25 | 19 | 8.8 |
| 1 ÷ 2.8 | 47 | 25 | 25.6 | 52 | 35 | 21.2 | 40 | 32 | 13.6 | 30 | 25 | 10.1 | 26 | 22 | 8.7 |

Note: For each range of $r_j$ the columns a,b and c are, respectively, the average number of nodes generated, the average number of nodes examined before optimum is reached, and the computer time in seconds.

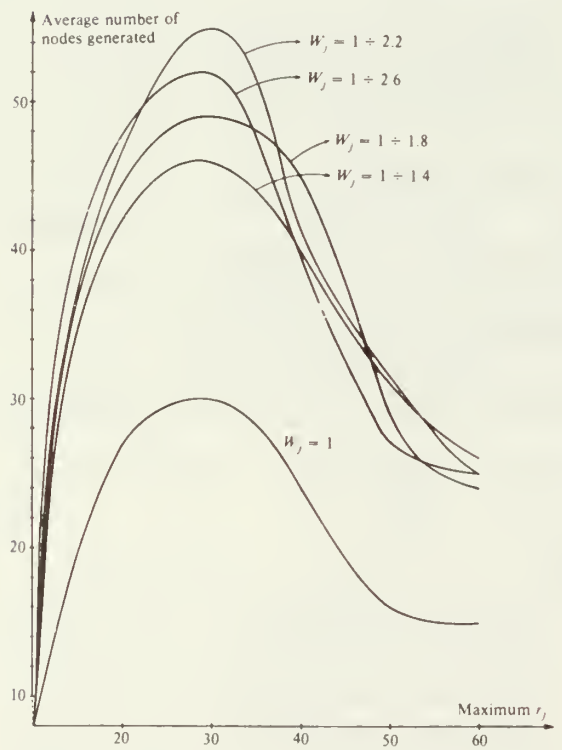



FIGURE 4 Nodes generated versus maximum ready time.

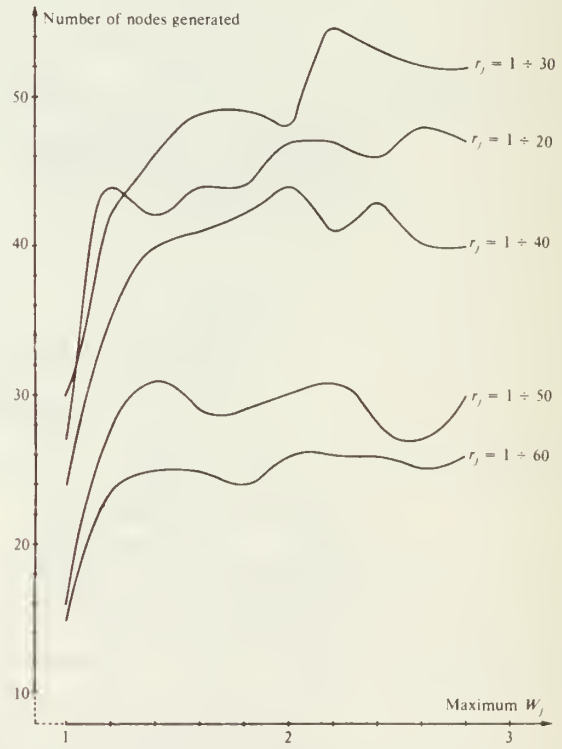



FIGURE 5. Nodes generated versus maximum weight.

In conclusion, the proposed algorithm appears to behave in a satisfactory way, at least when the number of jobs does not exceed 10.

However, to fully evaluate the algorithm performance, more extensive tests, by increasing the job number and the width of the weight ranges, by considering also processing time ranges of different width and by making use of a fast-time computer, should be provided.

## REFERENCES

[1] Baker, K.P. and Zan-Sing Su, "Sequencing with Due-Dates and Early Start Times to Minimize Maximum Tardiness," Naval Research Logistics Quarterly, *21*, 171-176 (1974).

[2] Bianco, L., B. Nicoletti, and S. Ricciardelli, "An Algorithm for Optimal Sequencing of Aircraft in the Near Terminal Area," Proceedings of the 8th IFIP Conference on Optimization Techniques (Springer Verlag, Berlin, 1977).

[3] Bianco, L. and S. Ricciardelli, "Il Problema del Sequenziamento Ottimo Degli Aerei in Area Terminale," Rapporto dell'Ist. di Automatica dell'Universitá di Roma e del C.S.S.C.C.A.-C.N.R. - R. 78-10 (1978).

[4] Bianco, L., S. Ricciardelli, A. Sassano, and G. Rinaldi, "The Aircraft Optimal Sequencing as an $N$ Job One Machine Scheduling Problem," EURO III — Amsterdam (1979) and Rapporto dell'Ist. di Automatica dell'Univ. di Roma e del C.S.S.C.C.A.-C.N.R., R. 79-36 (1979).

[5] Bratley, P., M. Florian, and P. Robillard, "On Sequencing with Earliest Starts and Due-Dates with Application to Computing Bounds for the $(n/m/G/F\max)$ Problem," Naval Research Logistics Quarterly, *20*, 57-67 (1973).

[6] Conway, R.W., W.L. Maxwell, and L.W. Miller, "Theory of Scheduling" (Addison-Wesley Publishing Co., Reading, MA, 1967).

[7] Emmons, H., "One Machine Sequencing to Minimize Certain Functions of Job Tardiness," Operations Research, *17*, 701-715 (1969).

[8] Fisher, M.L., "A Dual Algorithm for the One-Machine Scheduling Problem," Mathematical Programming, *11*, 229-251 (1976).

[9] Lenstra, J.K., A.H.G. Rinnoy Kan, and P. Brucker, "Complexity of Machine Scheduling Problems," Studies in Integer Programming. (North-Holland, Amsterdam, 1977).

[10] McMahon, G., and M. Florian, "On Scheduling with Ready Times and Due Dates to Minimize Maximum Lateness," Operations Reserch, *23*, 475-482 (1975).

[11] Rinaldi, G. and A. Sassano, "On a Job Scheduling Problem with different Ready Times: Some Properties and a New Algorithm to Determine the Optimal Solution," Rapporto dell'Ist. di Automatica dell'Universita di Roma e del C.S.S.C.C.A.-C.N.R,R., 77-24 (1977).

[12] Rinnoy Kan, A.H.G., B.J. Lageweg, and J.K. Lenstra, "Minimizing Total Costs in One-Machine Scheduling," Operations Research, *23*, 908-927 (1975).

[13] Shwimer, J., "On the $N$-Job, One-Machine, Sequence-Independent Scheduling Problem with Tardiness Penalties: a Branch-bound Solution," Management Science, *18*, B301-B313 (1970).

[14] Smith, W.E., "Various Optimizers for Single-Stage Production," Naval Research Logistics Quarterly, *3*, 59-66 (1956).

[15] Srinivasan, V., "A Hybrid Algorithm for the One-Machine Sequencing Problem to Minimize Total Tardiness," Naval Research Logistics Quarterly, *18*, 317-327 (1971).

## APPENDIX

**Proof of Theorem 1:**

Consider a schedule $s_n = s_k s_{n-k}$ having $h$ in position $k+1$ and $i$ in position $y > k+1$. Let $s'_n$ be a schedule that differs from $s_n$ only in that job $i$ is interchanged with its immediate predecessor $j$ in position $y-1$. The first $(y-2)$ jobs have exactly the same completion times under $s_n$ and $s'_n$ and, therefore make the same contribution to total weighted completion time.

Schedules $s_n$ and $s'_n$ differ only in the completion times of the $(n-y+2)$ remaining jobs.

As far as the last $(n-y)$ jobs are concerned, which are the same in both sequences, it can be derived that

$$(1) \qquad \Delta C^w(y) = \sum_{l=y+1}^{n} w_{[l]} C_{[l]}(s_n) - \sum_{l=y+1}^{n} w_{[l]} C_{[l]}(s'_n \geq 0.$$

In fact, by assumption b), it follows that

$$C_{[y]}(s_n) = \max(C_{[y-2]},\ r_j) + p_j + p_i$$
$$C_{[y]}(s'_n) = \max\{(C_{[y-2]} + p_i),\ r_j\} + p_j,$$

where $j$ is the job in position $y - 1$. Hence, it is easy to verify that

$$C_{[y]}(s_n) \geqslant C_{[y]}(s'_n)$$

from which the (1) immediately follows.

Let's examine now contribution of jobs $h$ and $i$. If we call $t = C_{[y-2]}(s_n)$, then their weighted completion times can be expressed by the following:

Under $s_n$:

$$w_j C_j = w_{[y-1]} C_{[y-1]} = w_j\{\max(t,\ r_j) + p_j\}$$
$$w_i C_i = w_{[y]} C_{[y]} = w_i\{\max(t,\ r_j) + p_j + p_i\};$$

Under $s'_n$:

$$w_j C_j = w_{[y]} C_{[y]} = w_j\{[\max(t + p_i),\ r_j] + p_j\}$$
$$w_i C_i = w_{[y-1]} C_{[y-1]} = w_i(t + p_i),$$

since, by assumptions b), $r_i \leqslant t_i \leqslant t_h < t$.

Two cases exist:

(1) $r_j \leqslant t$

Eliminating common terms, from the (1) and assumption (a), it follows that

$$C^w(s_n) - C^w(s'_n) = \Delta C^w(y) + w_i p_j - w_j p_i \geqslant 0$$

(2) $r_j > t$.

In this case

$$C^w(s_n) - C^w(s'_n) = \Delta C^w(y) + w_j(r_j + p_j) + w_i(r_i + p_i + p_i)$$
$$- w_j\{[\max(t + p_i),\ r_j] + p_j\} - w_i(t + p_i).$$

Two subcases are to be considered:

(2a) $r_j \geqslant t + p_i$.

From this and inequality (1) it follows that

$$C^w(s_n) - C^w(s'_n) = \Delta C^w(y) + w_i(r_j + p_j - t) > 0$$

(2b) $r_j < t + p_i$.

In this subcase, it is easy to verify that

$$C^w(s_n) - C^w(s'_n) = \Delta C^w(y) + (w_j + w_i)(r_j - t) + (w_i p_j - w_j p_i) > 0$$

since the first and third terms of the sum are nonnegative, by (1) and assumption (a), respectively, and the second term is positive, being $r_j > t$.

Hence, in conclusion, $s_n'$ represents an improvement over $s_n$. One could repeat the interchange operation on $s_n'$ by shifting job $i$ backward into position $y - 2$ and advancing the corresponding job into position $y - 1$. In this way the new sequence obtained is better than $s_n'$.

By induction, it is possible to repeat this operation on $s_n'$ until job $i$ reaches position $k + 2$. Then it follows that a partial sequence $s_k hi$ dominates every other partial sequence $s_k hj$, where $j \neq h$ and $i$ is an element of $\bar{K}$.

To complete the proof let's consider now as schedule $s_n$ the one with jobs $h$ and $i$ in the $(k + 1)$th and $(k + 2)$th positions, respectively.

Therefore $s_n'$ differs from $s_n$ in that job $i$ is in position $k + 1$ and job $h$ in position $k + 2$.

In this case, since $t = C_{[k]}(s_n)$ can be $< r_i$, only the completion times of $s_n'$ must be rewritten as follows:

Under $s_n'$:

$$w_h C_h = w_{[k+2]} C_{[k+2]} = w_h\{\max[\max(t, r_i) + p_i, r_h] + p_h\}$$
$$w_i C_i = w_{[k+1]} C_{[k+1]} = w_i\{\max(t, r_i) + p_i\}.$$

Three cases exist. Cases (1) and (2) have already been examined. Hence, one needs only to analyze the following:

(3) $t \leqslant r_i$.

From assumption (b) and definition of $t_j$, it is also true that $r_i \leqslant r_h$. Hence,

$$C^w(s_n) - C^w(s_n') = \Delta C^w(k + 2) + w_h(r_h + p_h) + w_i(r_h + p_h + p_i)$$
$$- w_h\{[\max(r_i + p_i), r_h] + p_h\} - w_i(r_i + p_i).$$

Two subcases are then to be examined:

(3a) $r_h \geqslant r_i + p_i$.

In this hypothesis it follows immediately that

$$C^w(s_n) - C^w(s_n') = \Delta C^w(k + 2) + w_i(r_h + p_h - r_i) > 0$$

(3b) $r_h < r_i + p_i$.

It is easy to verify that

$$C^w(s_n) - C^w(s_n') = \Delta C^w(k + 2) + (w_h + w_i)(r_h - r_i) + (w_i p_h - w_h p_i) \; )>= 0$$

since all terms of the sum are nonnegative by (1), assumptions (b) and (a), respectively.

In conclusion $s_n'$ is better than $s_n$.

Hence, since $s_k hi$ dominates every sequence with $s_k h$ fixed and $s_k ih$ always dominates $s_k hi$, it follows that $i$ dominates $h$ in position $k + 1$. Q.E.D.

Proof of Corollary 1.

It follows immediately from the previous theorem.

As far as theorems 2, 3, 4, 5 and Corollary 2 are concerned, they have been proved in [11].

Proof of Theorem 6:

It can be immediately seen that, for $k = n$, $\underline{C}^w(s_n) = C^w(s_n)$. Then it is only necessary to prove that for every $k$, $\underline{C}^w(s_n|s_k) \leqslant \underline{C}^w(s_n|s_{k+1})$.

$\underline{C}^w(s_n|s_k)$ and $\underline{C}^w(s_n|s_{k+1})$ can be expressed in the following way:

$$\underline{C}^w(s_n|s_{k+1}) = \underline{C}'(s_n|s_{k+1}) + \underline{C}''(s_n|s_{k+1})$$
$$\underline{C}^w(s_n|s_k) = \underline{C}'(s_n|s_k) + \underline{C}''(s_n|_k)$$

where

$$\underline{C}'(s_n|s_{k+1}) = C^w(s_k) + w_{k+1}C_{k+1}(s_k) + \sum_{j=1}^{n-k-1} w_{k+1+j}C_{k+1+j}(s_{k+1});$$

$$\underline{C}''(s_n|s_{k+1}) = \sum_{j=2}^{n-k-1}\sum_{l=j+1}^{n-k} \{\delta_{-1}[C_{k+j}(s_{k+1}) - t_{k+l}(s_{k+1})]$$

$$-\delta_{-1}[C_{k+j}(s_{k+1}) - C_{k+l}(s_{k+1})]\} \min(w_{k+j}, w_{k+l});$$

$$\underline{C}'(s_n|s_k) = C^w(s_k) + w_{k+1}C_{k+1}(s_k) + \sum_{j=1}^{n-k-1} w_{k+1+j}C_{k+1+j}(s_k)$$

$$+ \sum_{l=2}^{n-k} \{\delta_{-1}[C_{k+1}(s_k) - t_{k+l}(s_k)] - \delta_{-1}[C_{k+1}(s_k) - C_{k+l}(s_k)]\}.$$

$$\cdot \min(w_{k+1}, w_{k+l});$$

$$\underline{C}''(s_n|s_k) = \sum_{j=2}^{n-k-1}\sum_{l=j+1}^{n-k} \{\delta_{-1}[C_{k+j}(s_k) - t_{k+l}(s_k)]$$

$$- \delta_{-1}[C_{k+j}(s_k) - C_{k+l}(s_k)]\} \min(w_{k+j}, w_{k+l}).$$

Therefore, we will have:

$$\underline{C}'(s_n|s_{k+1}) - \underline{C}'(s_n|s_k) = \sum_{j=1}^{n-k-1} w_{k+1+j}[C_{k+1+j}(s_{k+1}) - C_{k+1+j}(s_k)]$$

$$- \sum_{l=2}^{n-k} \{\delta_{-1}[C_{k+1}(s_k) - t_{k+l}(s_k)] - \delta_{-1}[C_{k+1}(s_k) - C_{k+l}(s_k)]\}.$$

$$\cdot \min(w_{k+1}, w_{k+l}).$$

Recalling that $l = j+1$, the generic term of the difference can be expressed as follows

$$w_{k+l}[C_{k+l}(s_{k+1}) - C_{k+l}(s_k)] - \{\delta_{-1}[C_{k+1}(s_k) - t_{k+l}(s_k)] +$$
$$-\delta_{-1}[C_{k+1}(s_k) - C_{k+l}(s_k)]\} \min(w_{k+1}, w_{k+l}).$$

Let us consider the possible cases:

(a) $C_{k+1}(s_k) < t_{k+l}(s_k)$

(b) $t_{k+l}(s_k) \leqslant C_{k+1}(s_k) \leqslant C_{k+l}(s_k)$

(c) $C_{k+1}(s_k) > C_{k+l}(s_k)$

In case (a) it can be immediately verified that the generic term is equal to 0, while in case (b) and (c) it can be expressed, respectively, as follows:

$$\text{(b)}\quad w_{k+l}[C_{k+l}(s_{k+1}) - C_{k+l}(s_k)] - [C_{k+1}(s_k) - t_{k+l}(s_k)] \min(w_{k+1}, w_{k+l})$$
$$\geqslant 0, \text{ since } [C_{k+l}(s_{k+1}) - C_{k+l}(s_k)] = [C_{k+1}(s_k) - t_{k+l}(s_k)];$$
$$\text{(c)}\quad w_{k+l}[C_{k+l}(s_{k+1}) - C_{k+l}(s_k)] - p_{k+l} \cdot \min(w_{k+1}, w_{k+l}) \geqslant 0,$$
$$\text{since } [C_{k+l}(s_{k+1}) - C_{k+l}(s_k)] \geqslant p_{k+l}.$$

Then

$$\underline{C}'(s_n|s_{k+1}) - \underline{C}'(s_n|s_k) \geqslant 0.$$

Let us consider now the difference $\underline{C}''(s_n|s_{k+1}) - \underline{C}''(s_n|s_k)$. Calling the generic term $c_{k+1} - c_k$, we have:

$$c_{k+1} = \{\delta_{-1}[C_{k+j}(s_{k+1}) - t_{k+l}(s_{k+1})] - \delta_{-1}[C_{k+j}(s_{k+1})$$
$$-C_{k+l}(s_{k+1})]\} \min(w_{k+j}, w_{k+l});$$
$$c_k = \{\delta_{-1}[C_{k+j}(s_k) - t_{k+l}(s_k)] - \delta_{-1}[C_{k+j}(s_k) - C_{k+l}(s_k)].$$
$$\cdot \min(w_{k+j}, w_{k+l}).$$

The following cases can arise:

$$\text{(a)}\quad t_{k+l}(s_k) \leqslant C_{k+1}(s_k)$$
$$\text{(b)}\quad C_{k+1}(s_k) < t_{k+j}(s_k)$$
$$\text{(c)}\quad t_{k+j}(s_k) \leqslant C_{k+1}(s_k) < t_{k+l}(s_k).$$

Case a). According to the hypothesis

$$t_{k+j}(s_{k+1}) = t_{k+l}(s_{k+1}) = C_{k+1}(s_k)$$

and then

$$c_{k+1} = \{p_{k+j} - \delta_{-1}(p_{k+j} - p_{k+l})\} \min(w_{k+j}, w_{k+l})$$

Two subcases are possible:

$(a_1)$ If $p_{k+j} \leqslant p_{k+l}$, then

$$c_{k+1} = p_{k+j}\{\min(w_{k+j}, w_{k+l})\};$$
$$c_k = \begin{cases} 0 \\ [C_{k+j}(s_k) - t_{k+l}(s_k)]\min(w_{k+j}, w_{k+l}) \end{cases}.$$

Therefore, $c_{k+1} - c_k \geqslant 0$.

$(a_2)$ If $p_{k+j} > p_{k+l}$, then

$$c_{k+1} = p_{k+l}\{\min(w_{k+j}, w_{k+l})\}$$
$$c_k = \begin{cases} 0 \\ p_{k+l}\{\min(w_{k+j}, w_{k+l})\} \\ [C_{k+j}(s_k) - t_{k+l}(s_k)]\min(w_{k+j}, w_{k+l}) \end{cases}$$

Therefore, $c_{k+1} - c_k \geqslant 0$.

Case (b). According to the hypothesis we have

$$t_{k+j}(s_k) = t_{k+j}(s_{k+1}),\ t_{k+l}(s_k) = t_{k+l}(s_{k+1})$$

and then $c_{k+1} = c_k$.

Case (c). According to the hypothesis we have

$$t_{k+j}(s_{k+1}) = C_{k+1}(s_k) \text{ and } t_{k+l}(s_k) = t_{k+l}(s_{k+1}).$$

Three subcases exist.

$(c_1)$ $C_{k+j}(s_{k+1}) \leqslant t_{k+l}(s_{k+1})$ from which follows that

$c_{k+1} = c_k = 0.$

$(c_2)$ $t_{k+l}(s_{k+1}) < C_{k+j}(s_{k+1}) \leqslant C_{k+l}(s_{k+1}).$

In this case

$$c_{k+1} = [C_{k+j}(s_{k+1}) - t_{k+l}(s_{k+1})] \min(w_{k+j}, w_{k+l});$$

$$c_k = \begin{cases} 0 \\ [C_{k+j}(s_k) - t_{k+l}(s_k)] \min(w_{k+j}, w_{k+l}) \end{cases}$$

Therefore, $c_{k+1} - c_k \geqslant 0$.

$(c_3)$ $C_{k+l}(s_{k+1}) < C_{k+j}(s_{k+1}).$

In this case

$$c_{k+1} = p_{k+l}\{\min(w_{k+j}, w_{k+l})\}$$

$$c_k = \begin{cases} 0 \\ p_{k+l}\{\min(w_{k+j}, w_{k+l}) \\ [C_{k+j}(s_k) - t_{k+l}(s_k)] \min(w_{k+j}, w_{k+l}) \end{cases}$$

Therefore, $c_{k+1} - c_k \geqslant 0$.

Then, also, $\underline{C}''(s_n|s_{k+1}) - \underline{C}''(s_n|s_k) \geqslant 0$ and hence, in general, $\underline{C}^w(s_n|s_{k+1}) \geqslant \underline{C}^w(s_n|s_k)$, proving that $\underline{C}^w(s_n|s_k)$ is a lower bound for $(s_n|s_k)$.

Proof of Theorem 7.

First, we show that $C_{[k]}(s_n^0) \leqslant C_{[k]}(s_n^a)$, $1 \leqslant k \leqslant n$. To this end let's assume that the ready times are modified, defining $r_j' = \max(r_j, r_i)$, $\forall j \in N$.

This assumption will not affect the start or completion time of any job in $s_n^0$ or $s_n^a$.

Let's construct now a new sequence $s_n'$, reordering, for a given $k$, the first $k$ jobs in $s_n^a$ according to their earliest start times (EST) defined on the basis of ready times $r_j'$. If more jobs have the same EST, priority is given to the job with the smallest completion time among them.

Let's call the partial sequence so obtained $s_k'$ and the complete sequence $s_n'$; clearly $C_{[k]}(s_n') \leqslant C_{[k]}(s_n^a)$.

Since, by properties (a) and (b), $s_n^0$ is an ECT and EST sequence of the modified set $N$, it follows that, at the first position $l\,(1 \leqslant l \leqslant k)$ in which $C_{[l]}(s_n^0) \neq C_{[l]}(s_n')$, it must be true $C_{[l]}(s_n^0) < C_{[l]}(s_n')$ and $j_{[l]}^0 \notin s_k'$.

Hence, replacing $j_{[l]}'$ with $j_{[l]}^0$ will not result in an increase of $C_{[l']}(s_n')$, $l \leqslant l' \leqslant k$.

If this replacement causes jobs in $s_k'$ following the new $j_{[l]}'$ to lose the EST sequencing property, it is possible to reorder the set by EST without increasing $C_{[k]}(s_n')$, since $C_{[l]}(s_n')$ has been reduced. In this manner $s_k'$ is redefined to represent the new sequence. Repeating this procedure for all $l'\,(l \leqslant l' \leqslant k)$, comparing $C_{[l']}(s_n^0)$ and $C_{[l']}(s_n')$ and making necessary changes, as shown above, will produce a sequence identical to the partial sequence of the first $k$ jobs in $s_n^0$, without increasing $C_{[k]}(s_n^0)$. Thus, $C_{[k]}(s_n^0) \leqslant C_{[k]}(s_n') \leqslant C_{[k]}(s_n^a)$, $1 \leqslant k \leqslant n$, so proving the first statement. To complete the proof, remember that, by definition,

$$C^w(s_n^0) = \sum_{k=1}^{n} w_{[k]} C_{[k]}(s_n^0)$$

and

$$C^w(s_n^a) = \sum_{k=1}^{n} w_{[k]} C_{[k]}(s_n^a).$$

It is well known that such a sum of pairwise products of two sequences of numbers will be minimized if one sequence is arranged in increasing order and the other in decreasing order. Since the completion times $C_{[k]}$ are already in increasing order, the minimization of $C^w$ is accomplished if sequencing is such that the weights $w_{[k]}$ are in decreasing (or at least nonincreasing) order. Therefore, being $s_n^0$ and $s_n^a$ sequences defined on the same set $N$, from assumption (c) and previous results obtained, it follows that

$$C^w(s_n^0) \leqslant C^w(s_n^a)$$

which proves the theorem.

Proof of Corollary 3.

If follows immediately from the previous theorem.

# A NOTE ON SPLITTING THE BUMP IN AN ELIMINATION FACTORIZATION*

R. V. Helgason and J. L. Kennington

*Department of Operations Research*
*and Engineering Management*
*Southern Methodist University*
*Dallas, Texas*

**ABSTRACT**

This exposition presents a method for incorporating a technique known as "splitting the bump" within an elimination form reinversion algorithm. This procedure is designed to reduce fill-in during reinversion and should improve the efficiency of linear programming systems which already use the superior elimination form of the inverse.

## 1. INTRODUCTION

Current production linear programming systems are designed to handle problems with 8000 to 16,000 rows and 1000 row problems are considered to be medium sized (see [10, 17]). Fortunately, real problems tend to be sparse (the density of the constraint matrix is often less than 1%—see [1, 2, 9, 14, 15]). Hence, a basis for a 1000 row problem may have only (1000) (1000) (1%) = 10,000 nonzero entries. However, the inverse (which is required for the revised simplex algorithm) may be quite dense having almost a million nonzero elements. Consequently, one of the most important design considerations for a computer implementation of the revised simplex algorithm for general linear programs is the technique used to maintain and update the basis inverse.

In order to minimize the storage required to implement this algorithm, production linear programming systems maintain the inverse of the basis in either product or elimination form (see [1, 3, 8, 11, 12, 17]). Computationally, the inverse is stored as a sequence of vectors known as eta vectors, and the complete list of the eta's is called the ETA File. Each basis change (pivot) results in appending at least one eta vector to the ETA File. Since both the time per pivot and numerical error increase as the length of the ETA File increases, it is necessary to periodically obtain a new factorization of the basis inverse. This process of obtaining a new factorization is called reinversion. The objective of a reinversion algorithm is to obtain a factored inverse (i.e., ETA File) in which the sparsity characteristics of the original basis are preserved as much as possible.

The simplest reinversion technique for a given $m$-column basis can be thought of as successively reducing the basis to an identity matrix via $m$ pivot operations. The matrices which accomplish this reduction constitute the ETA File. Out of this simple approach, reinversion techniques have evolved which attempt to obtain a sparse factorization by selection of pivot positions, involving a reordering of the columns of the basis and a conceptual reordering of the rows. In addition, a technique known as splitting the bump, [8, 13], has been found beneficial when used within a product form reinversion

*This research was supported in part by the Air Force Office of Scientific Research under Grant Number AFOSR 77-3151.

algorithm. Recently, however, computational evidence indicates that the elimination form factorization is superior to the product form (with respect to both storage requirements and speed (see [2]). The objective of this exposition is to present a reinversion algorithm in which one may achieve some of the benefits of the "splitting the bump technique" when using the elimination form of the inverse.

For this note, the $i$th column of the $m \times m$ matrix B is denoted by $B(i)$. The symbol $e^i$ denotes the $m$-component column vector having $i$th component one, and all other components zero. The symbol $\eta$ denotes an $m$-component column vector and $\eta_i$ denotes the $i$th component of this vector.

## 2. FACTORIZATION ALGORITHMS

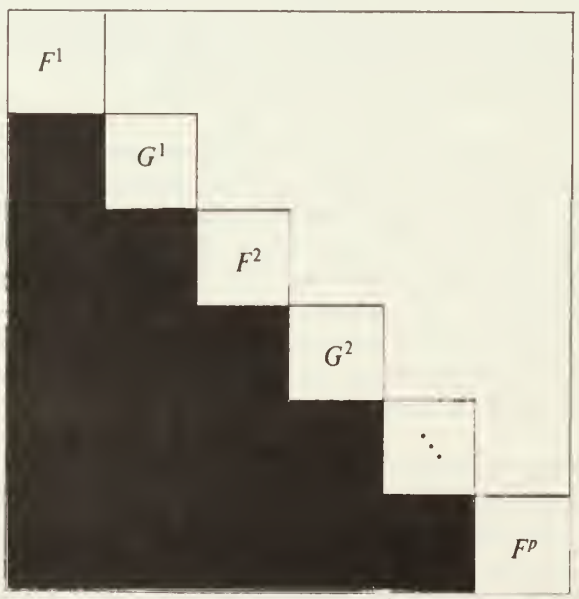
Let $B$ be any $m$ by $m$ nonsingular matrix. In this section we present two algorithms for obtaining a factorization of $B^{-1}$. The first algorithm produces a product form factorization which corresponds to the method for solving a system of linear equations known as Gauss-Jordan reduction while the second algorithm produces an elimination form factorization which corresponds to a Gauss reduction (see [4]).

By row and column interchanges, $B$ may be placed in the following form, where $B^1$ and $B^3$ are lower triangular matrices with nonzeroes on their diagonals:

(1)

| | | |
|---|---|---|
| $B^1$ | | |
| $B^4$ | $B^2$ | |
| $B^6$ | $B^5$ | $B^3$ |

We assume that if $B^2$ is nonvacuous, every row and column has at least two nonzero entries, so that no rearrangement of $B^2$ can expand the size of $B^1$ or $B^3$. $B^2$ is called the bump, merit, or heart section. We further require the heart section to assume the following form where $G^k$'s are either vacuous or lower triangular with nonzeroes on the diagonal




The only partitions in $B$ having columns with nonzeroes above the diagonal are the $F^k$'s which are called external bumps. The columns extending above the diagonal are called spikes or spike columns. An external bump is characterized as follows:

(i) the last column of an external bump will be a spike with a nonzero lying in the topmost row of the external bump, and

(ii) the nonspike columns have nonzero diagonal elements.

The algorithms of Hellerman and Rarick [6, 7] produce such a structure for any nonsingular matrix B, and we shall call a matrix having this structure an HR matrix.

The product form factorization algorithm for an HR matrix is given as follows:

ALGORITHM 1: PRODUCT FORM FACTORIZATION FOR A HR MATRIX

0. Initialization

Set $i \leftarrow 1$ and $\beta \leftarrow B(1)$.

1. Obtain New Eta

Set $$\eta_k \leftarrow \begin{cases} 1/\beta_i, & \text{for } k = i \\ -\beta_k/\beta_i, & \text{otherwise,} \end{cases}$$

and set $$E^i(k) \leftarrow \begin{cases} \eta, & \text{for } k = i \\ e^k, & \text{otherwise.} \end{cases}$$

(Note: only $\eta$ and $i$ need be saved in the ETA File rather than the matrix $E^i$)

2. Test For Termination

If $i = m$, terminate, otherwise, $i \leftarrow i + 1$.

3. Test For Spike

If $B(i)$ is not a spike, set $\beta \leftarrow B(i)$ and go to 1.

4. Spike Update

Let $B(k)$ correspond to the first column of the external bump containing $B(i)$. Set $\beta \leftarrow E^{i-1} \ldots E^k B(i)$.

5. Swap Spikes If Pivot Element Equals Zero

If $\beta_i \neq 0$, go to 1; otherwise, there is some spike $B(j)$ in the same external bump having $j > i$ such that the $i$th component of $E^{i-1} \ldots E^k B(j)$ is nonzero. Set $\beta \leftarrow E^{i-1} \ldots E^k B(j)$, interchange $B(i)$ and $B(j)$ and go to 1.

In practice, the test for a zero pivot element in Step 5 is usually replaced by a tolerance test. Let TOL denote the tolerance to be used in the test. If $|\beta_i| \leqslant$ TOL, then $\beta_i$ is treated as if it is zero. Discussions of tolerances may be found in [12, 15, 16]. Similar tolerance tests would ordinarily be incorporated in the other algorithms to be presented in this exposition. To simplify the presentation they have been omitted here.

Justification for Algorithm 1 is given in [5]. At termination, $B^{-1} = E^m E^{m-1} \ldots E^1$. Furthermore, we see that for nonspike columns, each $\beta$ and consequently each $\eta$ has exactly the same number of nonzeroes as the corresponding column of $B$. However, the $\eta$ for a spike column may have a higher density than the original column of $B$. The phenomena of an $\eta$ having a higher density than the corresponding column of $B$ is known as fill-in.

It is well known that fill-in may be reduced by applying an elimination form factorization rather than the product form (see [1]). The elimination form algorithm for an HR matrix is given as follows:

ALGORITHM 2: ELIMINATION FORM FACTORIZATION FOR A HR MATRIX

0. Initialization

Set $i \leftarrow 1$, $r \leftarrow 0$, and $\beta \leftarrow B(1)$.

1. Obtain New Lower Eta

$$\text{Set} \quad \eta_k \leftarrow \begin{cases} 1/\beta_i, & \text{for } k = i \\ -\beta_k/\beta_i, & \text{for } k > i \\ 0, & \text{otherwise,} \end{cases}$$

$$\text{and set } L^i(k) \leftarrow \begin{cases} \eta, & \text{for } k = i \\ e^k, & \text{otherwise.} \end{cases}$$

(Note: only $\eta$ and $i$ need be saved in the ETA File for the lower factors)

2. Test for Termination

If $i = m$, terminate; otherwise, $i \leftarrow i + 1$.

3. Test for Spike

If $B(i)$ is not a spike, set $\beta \leftarrow B(i)$ and go to 1.

4. Spike Update

Let $B(k)$ correspond to the first column of the external bump containing $B(i)$. Set $\beta \leftarrow L^{i-1} \ldots L^k B(i)$.

5. Swap Spikes if Pivot Element Equals Zero

If $\beta_i \neq 0$, go to 6; otherwise, there is some spike $B(j)$ in the same external bump having $j > i$ such that the $i$th component of $L^{i-1} \ldots L^k B(j)$ is nonzero. Set $\beta \leftarrow L^{i-1} \ldots L^k B(j)$ and interchange $B(i)$ and $B(j)$.

6. Obtain New Upper Eta

Set $r \leftarrow r + 1$

$$\text{Set} \quad \eta_k \leftarrow \begin{cases} 1, & \text{for } k = i \\ -\beta_k, & \text{for } k < i \\ 0, & \text{otherwise,} \end{cases}$$

$$\text{set} \quad U^r(k) \leftarrow \begin{cases} \eta, & \text{for } k = i \\ e^k, & \text{otherwise,} \end{cases}$$

and go to 1.

(Note: only $\eta$ and $i$ need be saved in the ETA File for the upper factors). At the termination of Algorithm 2, $B^{-1} = U_1 \ldots U^r L^m \ldots L^1$, where the upper eta's are upper triangular and the lower eta's are lower triangular. As with Algorithm 1 fill-in has been restricted to the eta's corresponding to spike columns.

## 3. SPLITTING THE BUMP

In an attempt to reduce the fill-in which occurs in spike columns during a reinversion using Algorithm 1, a variation of the elimination form factorization algorithm (attributed to Martin Beale, see [13]) has been used by some practitioners. This technique has been called "splitting the bump" after its treatment of external bump columns.

Consider a set of columns of the basis corresponding to an external bump, say

$$B^3 = \begin{bmatrix} 0 \\ --- \\ F \\ --- \\ H \end{bmatrix} \begin{matrix} \\ \\ \} \ n \text{ rows} \\ \\ \} \ p \text{ rows} \end{matrix}$$

where $F$ is an external bump. Suppose $F$ contains $q$ spikes. If we apply the standard product form algorithm, we obtain a set of eta's such that

$$E^n \ldots E^1 B^3 = \begin{bmatrix} 0 \\ --- \\ I \\ --- \\ 0 \end{bmatrix} \begin{matrix} \\ \\ \} n \\ \\ \} p \end{matrix} .$$

The eta's corresponding to the $q$ spikes may incur fill-in. If we split the bump, the fill-in can be restricted to the $n$ rows corresponding to $F$; hence, we may avoid fill-in in the last $p$ rows. Since $p$ may be much larger than $n$, the savings could be substantial. The price which must be paid to attain this reduction in fill-in is that each external bump will require $2n + q$ eta's rather than $n$. The savings in fill-in is typically so great that it offsets the additional storage which must be given up for the additional eta's.

The product form algorithm incorporating the "splitting the bump" technique is given as follows:

ALGORITHM 3: PRODUCT FORM FACTORIZATION FOR A HR MATRIX INCLUDING SPLITTING THE BUMP

0. Initialization

Set $i \leftarrow 1$, $j \leftarrow 1$, and $\beta \leftarrow B(1)$. If $B(1)$ is in an external bump, go to 4.

1. Obtain New Lower Eta

Set $$\eta_k \leftarrow \begin{cases} 1/\beta_i, & \text{for } k = i \\ -\beta_k/\beta_i, & \text{otherwise,} \end{cases}$$

and set $$E^j(k) \leftarrow \begin{cases} \eta, & \text{for } k = i \\ e^k, & \text{otherwise.} \end{cases}$$

Set $j \leftarrow j + 1$.

2. Test for Termination

If $i = m$, terminate; otherwise, $i \leftarrow i + 1$ and $\beta \leftarrow B(i)$.

3. Test for External Bump

If $B(i)$ is not the first column in an external bump, go to 1.

4. Initialization for Bump

Set $\begin{cases} l \leftarrow j \text{ (current length of Eta File)} \\ s \leftarrow i \text{ (first column in this external bump)} \\ b \leftarrow \text{number of columns in this external bump} \\ t \leftarrow i + b - 1 \text{ (last column in this external bump)} \end{cases}$

5. Obtain Lower Eta

Set $\eta_k \leftarrow \begin{cases} 1/\beta_i, & \text{for } k = i \\ -\beta_k/\beta_i, & \text{for } i < k \leqslant t \\ 0, & \text{otherwise,} \end{cases}$

and set $E^j(k) \leftarrow \begin{cases} \eta, & \text{for } k = i \\ e^k, & \text{otherwise.} \end{cases}$

Set $j \leftarrow j + 1$.

6. Test for End of Bump

If $i = t$, go to 10; otherwise, $i \leftarrow i + 1$.

7. Test for Spike

If $B(i)$ is not a spike, set $\beta \leftarrow B(i)$ and go to 5.

8. Spike Update

Set $\beta \leftarrow E^{j-1} \ldots E^1 B(i)$.

9. Swap Spikes if Pivot Element Equals Zero

If $\beta_i \neq 0$, go to 5; otherwise, there is some spike $B(r)$ in the same external bump having $r > i$ such that the $i$th component of $E^{j-1} \ldots E^1 B(r)$ is nonzero. Set $\beta \leftarrow E^{j-1} \ldots E^1 B(r)$, interchange $B(r)$ and $B(i)$ and go to 5.

10. Obtain Upper Eta

Set $\eta_k \leftarrow \begin{cases} 1, & \text{for } k = i \\ -\beta_k, & \text{for } k < i \\ 0, & \text{otherwise,} \end{cases}$

and set $E^j(k) \leftarrow \begin{cases} \eta, & \text{for } k = i \\ e^k, & \text{otherwise.} \end{cases}$

Set $j \leftarrow +1$ and $i \leftarrow i - 1$.

11. Test for Beginning of Bump

If $i = s$, set $\beta \leftarrow B(i)$ and go to 13.

12. Test for Spike

If $B(i)$ is not a spike, set $i \leftarrow i - 1$ and go to 11; otherwise, set $\beta \leftarrow E^{j-1} \ldots E^{l}B(i)$ and go to 10.

13. Obtain Lower Eta

$$\text{Set} \quad \eta_k \leftarrow \begin{cases} 1, \text{ for } k = i \\ -\beta_k, \text{ for } k > t \\ 0, \text{ otherwise,} \end{cases}$$

$$\text{and set } E^j(k) \leftarrow \begin{cases} \eta, \text{ for } k = i \\ e^k, \text{ otherwise.} \end{cases}$$

Set $j \leftarrow j + 1$.

14. Test for End of Bump

If $i = t$, go to 2; otherwise $i \leftarrow i + 1$, set $\beta \leftarrow B(i)$, and go to 13.

At the termination of Algorithm 3, $B^{-1} = E^{j-1} \ldots E^1$. Furthermore, all eta's are either upper triangular or lower triangular, but they are intermixed.

Recall that Algorithm 2 produces a set of lower triangular factors followed by a set of upper triangular factors. If one incorporates the "splitting the bump" technique into Algorithm 2 and applies it to an HR matrix, the lower and upper factors become intermingled. Once the factors have become intermingled, we may no longer use the important algorithm of Forrest and Tomlin [2] to maintain the elimination form. We now show how one may achieve a partial "splitting of the bump" while simultaneously maintaining a partitioning of the upper and lower factors.

Recall that an HR matrix takes the form given in (1).

By a rearrangement of rows and columns, the HR matrix (1) may be placed in the following form:

$$\bar{B} = \begin{array}{|c|c|c|} \hline \bar{B}^3 & \bar{B}^6 & \bar{B}^5 \\ \hline & B^1 & \\ \hline & B^4 & B^2 \\ \hline \end{array}$$

$$\underbrace{\phantom{xx}}_{u} \quad \underbrace{\phantom{xx}}_{v} \quad \underbrace{\phantom{xx}}_{w}$$

$\Rightarrow$

where $\bar{B}^6$ and $\bar{B}^5$ are row permutations of $B^6$ and $B^5$, respectively, and $\bar{B}^3$ is a row and column permutation of $B^3$. Applying a variation of Algorithm 2 to $\bar{B}$ eliminates all fill-in in $\bar{B}^5$ while representing $\bar{B}^{-1}$ as a product of upper factors followed by a product of lower factors. The details are given in the following algorithm.

## ALGORITHM 4: ELIMINATION FORM FACTORIZATION FOR A MODIFIED HR MATRIX

0. Initialization

Set $i \leftarrow u + 1$ and $\beta \leftarrow \bar{B}(i)$. If v = 0, go to 3.

1. Obtain New Lower Eta

$$\text{Set} \quad \eta_k \leftarrow \begin{cases} 1/\beta_i, & \text{for } k = i \\ -\beta_k/\beta_i, & \text{for } k > i \\ 0, & \text{otherwise,} \end{cases}$$

$$\text{and set} \quad L^i(k) \leftarrow \begin{cases} \eta, & \text{for } k = i \\ e^k, & \text{otherwise.} \end{cases}$$

2. Test for End of Section 2

Set $i \leftarrow i + 1$ and $\beta \leftarrow \bar{B}(i)$. If $i \leqslant u + \text{v}$, go to 1. If $w = 0$, go to 8.

3. Obtain New Lower Eta

$$\text{Set} \quad \eta_k \leftarrow \begin{cases} 1/\beta_i, & \text{for } k = i \\ -\beta_k/\beta_i, & \text{for } k > i \\ 0, & \text{otherwise,} \end{cases}$$

$$\text{and set} \quad L^i(k) \leftarrow \begin{cases} \eta, & \text{for } k = i \\ e^k, & \text{otherwise.} \end{cases}$$

4. Test for End of Section 3

If $i = m$, go to 8; otherwise, $i \leftarrow i + 1$.

5. Test for Spike

If $\bar{B}(i)$ is not a spike, set $\beta \leftarrow \bar{B}(i)$ and go to 3.

6. Spike Update

Set $\beta \leftarrow L^{i-1} \ldots L^{u+\text{v}+1} \bar{B}(i)$.

7. Swap Spikes if Pivot Element Equals Zero

If $\beta_i \neq 0$, go to 3; otherwise, there is some spike $B(j)$ in the same external bump having $j > i$ such that the $i$th component of $L^{i-1} \ldots L^{u+\text{v}+1} \bar{B}(j)$ is nonzero. Set $\beta \leftarrow L^{i-1} \ldots L^{u+\text{v}+1} \bar{B}(j)$, interchange $\bar{B}(i)$ and $\bar{B}(j)$ and go to 3.

8. Obtain New Upper Eta

$$\text{Set} \quad \eta_k \leftarrow \begin{cases} 1, \text{ for } k = i \\ -\beta_k, \text{ for } k < i \\ 0, \text{ otherwise,} \end{cases}$$

$$\text{and set} \quad U^i(k) \leftarrow \begin{cases} \eta, \text{ for } k = i \\ e^k, \text{ otherwise.} \end{cases}$$

9. Check for End of Section 2

Set $i \leftarrow i - 1$. If $i = 0$, terminate.
If $i = u$, set $\beta \leftarrow B(i)$ and go to 12.

10. Set Column

If $i > u + r$, set $\beta \leftarrow L^i \ldots L^{u+v+1} \bar{B}(i)$; otherwise,
set $\beta \leftarrow \bar{B}(i)$. Go to 8.

11. Obtain New Upper Eta

$$\text{Set} \quad \eta_k \leftarrow \begin{cases} 1/\beta_i, \text{ for } k = i \\ -\beta_k/\beta_i, \text{ for } k < i \\ 0, \text{ otherwise,} \end{cases}$$

$$\text{and set} \quad U^i(k) \leftarrow \begin{cases} \eta, \text{ for } k = i \\ e^k, \text{ otherwise.} \end{cases}$$

12. Check for Termination

If $i = 1$, terminate; otherwise $\beta \leftarrow \bar{B}(i)$ and go to 11.

At the termination of Algorithm 4, $\bar{B}^{-1} = U^1 \ldots U^m L^m \ldots L^{u+1}$, and the fill-in has been restricted to $B^2$. Hence, Algorithm 4 gains some of the benefits of the "splitting the bump" technique while maintaining a partitioning of the upper and lower factors. The benefits are not as great as with ordinary "bump splitting" since each individual external bump is split, whereas here the split is with respect to the entire heart section.

A variation of Algorithm 4 has also been used by Tomlin [15]. Our contribution is that we have tied together the ideas of "splitting the bump" in both the product and elimination form factorizations and we have explicitly indicated via Algorithms 3 and 4 how these may be implemented.

## REFERENCES

[1] Brayton, R.K., F.G. Gustavson and R.A. Willoughby, "Some Results on Sparse Matrices," Mathematics of Computation, *24*, 112, 937-954 (1970).
[2] Forrest, J.J.H. and J.A. Tomlin, "Updated Triangular Factors of the Basis to Maintain Sparsity in the Product from Simplex Method," Mathematical Programming, *2*, 263-278 (1972).
[3] Greenberg, H.J. and D.C. Rarick, "Determining GUB Sets Via an Invert Agenda Algorithm," Mathematical Programming, *7*, 240-244 (1974).
[4] Hadley, G., *Linear Algebra* (Addison Wesley Publishing Co., Inc., Reading, MA, 1964).

[5] Helgason, R.V. and J.L. Kennington, "Spike Swapping in Basis Reinversion," Naval Research Logistics Quarterly, *27*, 697-701 (1980).
[6] Hellerman, E. and D. Rarick, "Reinversion With the Preassigned Pivot Procedure," Mathematical Programming, *1*, 195-216 (1971).
[7] Hellerman, E. and D. Rarick, "The Partitioned Preassigned Pivot Procedure ($P^4$)," Sparse Matrices and Their Applications, Edited by D. Rose and R. Willoughby (Plenum Press, New York, 1972).
[8] Kalan, J.E., "Aspects of Large-Scale In-Core Linear Programming," Proceedings of the ACM, 304-313 (1971).
[9] Luhks, R.A., "Large-Scale Linear Programming Problems: Analysis and Generation," unpublished dissertation, Department of Computer Science and Operations Research, Southern Methodist University, Dallas, TX (1972).
[10] Marceau, I.W., R.J. Lermit, J.C. McMillan, T. Yamamoto and S.S. Yardley, "Linear Programming on Illiac IV," Proceedings of Fourth Austrailian Computer Conference, Adelaide, Australia, 481-485 (1969).
[11] Markowitz, H.M., "The Elimination Form of the Inverse and Its Application to Linear Programming," Management Science, *3*, 255-269 (1975).
[12] Orchard-Hays, W., *Advanced Linear-Programming Computing Techniques,* (McGraw-Hill, New York, 1968).
[13] Orchard-Hays, W., "An Exercise in Basis Inversion," SIGMAP Newsletter, No. 15 (1973).
[14] Tomlin, J.A., "Modifying Triangular Factors of the Basis in the Simplex Method," in *Sparse Matrices and Their Applications*, Edited by D.J. Rose and R.A. Willoughby (Plenum Press, New York, 1972).
[15] Tomlin, J.A., "Pivoting for Size and Sparsity in Linear Programming Inversion Routines," Journal of the Institute of Mathematics and its Applications, *10*, 289-295 (1972).
[16] Tomlin, J.A., "An Accuracy Test for Updating Triangular Factors," *Mathematical Programing Study 4*, 142-145 (1975).
[17] White, W.W., "A Status Report on Computing Algorithms for Mathematical Programming," Computing Surveys, *5*, 135-166 (1973).

# A NOTE ON SOLVING MULTIFACILITY LOCATION PROBLEMS INVOLVING EUCLIDEAN DISTANCES

Henrik Juel

*Soenderborg School of Economics and Business Administration*
*Soenderborg, Denmark*

ABSTRACT

This note considers a recently proposed solution method for a multifacility location problem. It is shown that the method does not always produce an optimal solution.

In a recent issue of this journal Calamai and Charalambous [1] conclude that many existing methods designed for solving the multifacility location problem are either poorly structured, suboptimal, or haphazard (p. 619). It should be pointed out that their proposed algorithm belongs to the same class of methods, in the sense that the algorithm sometimes gives a suboptimal solution.

The authors base their algorithm on a set of necessary conditions and seem to believe that this set of conditions is also sufficient for optimality, although sufficiency has yet to be proven (p. 617). In a problem involving rectilinear distances the necessary conditions are also sufficient [3]. This result holds for Euclidean distances, too, as long as each case of facility coincidence involves just two facilities [2]. But if a current solution to a problem involving Euclidean distances has three or more facilities located at the same point, the author's set of necessary conditions are not in general sufficient for optimality, as shown by the following example.

Consider a current solution where new facilities 1, 2, and 3 occupy the same location and no other facility occupies this location. Let the given weights relating to these new facilities be $v_{12} = 1$, $v_{13} = 1$, and $v_{23} = 5^{1/2} - 1$. Suppose the other weights and facility locations are such that the pseudo-gradients in (4.2c) (p. 613) are $G_1 = (2, 0)$, $G_2 = (0, 1)$, and $G_3 = (-2, -1)$. Now consider the necessary conditions for optimality in (5.3) (p. 613). When $S$ contains one element, the left hand sides of the condition are 2, 1, and $5^{1/2}$, whereas the right hand sides are 2, $5^{1/2}$, and $5^{1/2}$. Hence, moving one of the new facilities 1, 2, or 3, as done in step (3) (p. 615) of the authors' algorithm, does not yield a better solution. Moving the entire cluster of these new facilities is likewise unsuccessful, since the sum of the pseudo-gradients is (0, 0). Finally, when $S$ contains two elements, the conditions for $S$ containing one element are repeated so moving a subset cluster, as done in step (5) of the algorithm, would not yield a better solution.

Summarizing, the example represents a situation where the algorithm will stop, because the necessary conditions for optimality are satisfied. But the current solution is suboptimal, because the directional derivative is not nonnegative for all directions. To verify this consider, for instance, a direction vector $D$ where the only nonvanishing two-dimensional component vectors are $D_1 = (-4, -1)$ and $D = (-2, -1)$. ($D$ can be normalized into a unit vector by using the factor $22^{1/2}$, if desired.) The directional derivative at the current solution in this direction is given by (4.1) (p. 612), which yields: $-8-1+0+2+17^{1/2}+(5-5^{1/2})$, which is negative. Thus, a better solution can be achieved by keeping new facility 3 fixed, and by moving new facilities 1 and 2, not as a subcluster, but along the separate directions indicated above.

The possibility that facilities may coincide renders the objective function in the multifacility location problem nondifferentiable, and, thus, difficult to handle computationally. Until a satisfactory nondifferentiable optimization technique has been developed, decision makers should continue to use a differentiable approximation to their problem (unless the distances are rectilinear). The new algorithm is of dubious value to a decision maker interested in an optimal solution to the problem.

Finally, it might be pointed out that problem 6 (p. 618) can be solved by inspection using corollary 2 and corollary 1 in [4]. The results in this reference offer the potential for reduction of practical problems by exploiting the special weight structure.

## ACKNOWLEDGMENT

The author wishes to thank an anonymous referee, whose comments made significant improvements possible in the original manuscript.

## REFERENCES

[1] Calamai, P. and C. Charalambous, "Solving Multifacility Location Problems Involving Euclidean Distances," Naval Research Logistics Quarterly, *27*, 609-620 (1980).
[2] Juel, H., "Solution Conditions for Generalized Weber Problems with Two New Facilities," Transportation Science (to appear).
[3] Juel, H. and R.F. Love, "An Efficient Computational Procedure for Solving the Multi-Facility Rectilinear Facilities Location Problem," Operational Research Quarterly, *27*, 697-703 (1976).
[4] Juel, H. and R.F. Love, "Sufficient Conditions for Optimal Facility Locations to Coincide," Transportation Science, *14*, 125-129 (1980).

Readers who wish to subscribe to the Naval Research Logistics Quarterly may use this form. The cost for a one year subscription is $15.00 in the United States and Canada, $18.75 elsewhere. Subscription orders should be sent to Superintendent of Documents, U.S. Government Printing Office, Washington, D.C. 20402. Department of Defense activities may request distribution by writing to Office of Naval Research, Arlington, VA 22217.

**SUBSCRIPTION ORDER FORM**

Deposit Account Number

| | | | | | | | – | |
|---|---|---|---|---|---|---|---|---|
| | | | | | | | | |

*Enclosed find $_____ (check or money order), or charge to my Superintendent of Documents Deposit Account Number shown above.*

*Your Order Number* _____

*Please send the subscription(s) I have listed to the address shown at the bottom of this form.*

MAIL ORDER FORM TO:
Superintendent of Documents
U.S. Government Printing Office
Washington, D.C. 20402

| FOR OFFICE USE ONLY | |
|---|---|
| Quantity | Charges |
| ------ Enclosed | |
| ------ To be mailed | |
| ------ Subscriptions | |
| Postage | |
| Foreign handling | |
| MMOB | |
| OPNR | |
| ------ UPNS | |
| ------ Discount | |
| ------ Refund | |

MasterCard

VISA

Credit Card Orders Only
Total charges $_____. Fill in the boxes below:

Credit Card No. | | | | | | | | | | | | | | | | |

Expiration Date Month/Year | | | | |

*NOTE: Complete top address portion if different from that at the bottom.*

Name _____

Street address _____

City and State _____ ZIP Code _____

| SYMBOL | TITLE | PRICE |
|---|---|---|
| NRLQ | Naval Research Logistics Quarterly | |
| | | |
| | | |
| | | |
| | | |

COMPANY OR PERSONAL NAME

ADDITIONAL ADDRESS/ATTENTION LINE

STREET ADDRESS

CITY | STATE | ZIP CODE

(or) COUNTRY

PLEASE PRINT

## INFORMATION FOR CONTRIBUTORS

The NAVAL RESEARCH LOGISTICS QUARTERLY is devoted to the dissemination of scientific information in logistics and will publish research and expository papers, including those in certain areas of mathematics, statistics, and economics, relevant to the over-all effort to improve the efficiency and effectiveness of logistics operations.

Manuscripts and other items for publication should be sent to The Managing Editor, NAVAL RESEARCH LOGISTICS QUARTERLY, Office of Naval Research, Arlington, Va. 22217. Each manuscript which is considered to be suitable material for the QUARTERLY is sent to one or more referees.

Manuscripts submitted for publication should be typewritten, double-spaced, and the author should retain a copy. Refereeing may be expedited if an extra copy of the manuscript is submitted with the original.

A short abstract (not over 400 words) should accompany each manuscript. This will appear at the head of the published paper in the QUARTERLY.

There is no authorization for compensation to authors for papers which have been accepted for publication. Authors will receive 250 reprints of their published papers.

Readers are invited to submit to the Managing Editor items of general interest in the field of logistics, for possible publication in the NEWS AND MEMORANDA or NOTES sections of the QUARTERLY.

NAVAL RESEARCH
LOGISTICS
QUARTERLY

MARCH 1982
VOL. 29, NO. 1
NAVSO P-1278

## CONTENTS


| ARTICLES | | Page |
|---|---|---|
| A Theory of Multistage Contractual Incentives with Application to Design-to-Cost | R.W. BLANNING<br>P.R. KLEINDORFER<br>C.S. SANKAR | 1 |
| Optimal Strategies in a Game of Economic Survival | K. BORCH | 19 |
| Diagnostic Analysis of Inventory Systems: an Optimization Approach | G.R. BITRAN<br>A.C. HAX<br>J. VALOR-SABATIER | 29 |
| Production Planning for Multi-Resource Network Systems | R.C. LEACHMAN | 47 |
| The Class of MIFRA Lifetimes and Its Relation to Other Classes | H.W. BLOCK<br>T.H. SAVITS | 55 |
| Remarks on a Univariate Shock Model with Some Bivariate Generalizations | W.S. GRIFFITH | 63 |
| Asymptotic Joint Normality of an Increasing Number of Multivariate Order Statistics and Associated Cell Frequencies | L. WEISS | 75 |
| Minimizing Costs of Sequential and Parallel Personnel Testing Programs | P.D. FEIGIN<br>G.H. HAREL | 87 |
| The Variance of the Coverage of a Randomly Located Area Target by a Salvo of Weapons | G. SCHROETER | 97 |
| Reformulating Zero-Sum Games with Multiple Goals | E.L. HANNAN | 113 |
| On a Generalized Bin-Packing Problem | R.T. LEWIS<br>R.G. PARKER | 119 |
| An Integer Programming Algorithm for Portfolio Selection with Fixed Charges | M.W. COOPER<br>K. FARHANGIAN | 147 |
| Scheduling of a Single Machine to Minimize Total Weighted Completion Time Subject to Release Dates | L. BIANCO<br>S. RICCIARDELLI | 151 |
| A Note on Splitting the Bump in an Elimination Factorization | R.V. HELGASON<br>J.L. KENNINGTON | 169 |
| A Note on Solving Multifacility Location Problems Involving Euclidean Distances | H. JUEL | 179 |


OFFICE OF NAVAL RESEARCH
Arlington, Va. 22217